Genomic analysis and metabolic modelling of Geobacillus thermoglucosidasiusNCIMB 11955

Lisowska, Beata

Award date:
2016

Awarding institution:
University of Bath

Genomic analysis and metabolic modelling of *Geobacillus thermoglucosidasius* NCIMB 11955

University of Bath

Department of Biology and Biochemistry

Thesis submitted for the degree of Doctor of Philosophy of University of Bath

Beata Karolina Lisowska

January 2016

The work in this thesis is my own and has not been submitted previously for any award. All other sources of information used in this document have been acknowledged appropriately.

Declaration of material from previously submitted thesis and of work done in conjunction with others

PathwayBooster was jointly developed with Dr Rodrigo Liberal at Imperial College London. The work division is clearly defined in the introduction to Chapter "PathwayBooster".

The experimental estimations of biomass components were done by Dr Shyam Masakapalli. Dr Shyam Masakapalli and Dr Leann Bacon conducted RNA extraction for the RNAseq analysis that was done at Deep Seq: Next Generation Sequencing Facility at the University of Nottingham. Dr Alice Marriott and Dr Shyam Masakapalli conducted chemostat experiments in sections 6.2.1, 6.3.1, 6.4.1 in conjunction with the author.

Dr Steven Bowden and Carolyn Williamson took the electron microscope image of *G. thermoglucosidasius* NCIMB 11955 microcompartments (section 3.4).

Some GAPDH cloning procedures were firstly carried out by Andrew Balfour under my supervision. These can be found in section 3.2.2 where the student is credited.

Abstract

Geobacillus thermoglucosidasius is a Gram-positive thermophilic eubacterium (45-70℃) that has the ability to convert pre-treated lignocellulosic material LCM into ethanol. This organism has been genetically engineered such that its yield of ethanol production is in excess of 90% of the theoretical maximum [38]. There remains considerable scope to develop *G. thermoglucosidasius* to produce alternative fuels and chemicals of industrial importance.

For such a useful bacterium the understanding of the global metabolism remains poorly characterised. To gain a better insight into the metabolic pathways and capabilities of *G. thermoglucosidasius* a bottom-up approach to construct a comprehensive metabolic model of the organism was applied. The model was build from manually annotated genome and incorporates data from wet lab experiments for accurate *in silico* analyses. The model simulations has highlighted a potential experimental design for the *in silico* production of succinate and butane-2,3-diol.

PathwayBooster is also introduced in this study as a tool for curating metabolic pathways. The methodology is based on the assumption that the core metabolic capabilities are shared among evolutionarily closely related species [80]. This approach led to the further analysis of members of the genus *Geobacillus* with respect to their core metabolic capabilities, genome re-arrangements and shared unique features. Theoretical route for the biosynthesis of Vitamin B_{12} is presented here, which is novel to the canonical aerobic and anaerobic pathways known to date and ubiquitous amongst *Geobacillus* spp.

The analysis of the gene assignment for this bacterium has highlighted the presence of NADP-dependent GAPDH. The theoretical function of this novel and previously uncategorised enzyme in the genus *Geobacillus* has been confirmed through enzymatic assays.

Acknowledgements

I would like to thank Professor David Leak for being such a patient and supportive mentor over the past four years. He has guided me through troubled waters both in academic and personal capacity and for that I will be eternally grateful to him. I would like to thank Professor Michael Danson for sharing his knowledge and kindness, whenever and whichever was needed at the time. I also owe my deepest gratitude to Dr John Pinney at Imperial College London for his support, even after my move to the University of Bath.

I would like to thank everyone at the Leak and Danson Labs for their friendship and motivation when I was tired, frustrated and grumpy, including: Giannina Espina, Carolyn Williamson, Chris Vennard, Dr Steve Bowden, Dr Steven Bowden, Dr Jeremy Bartosiak-Jentys, Lisa Buddrus, Dr Matthew Styles, Dr Christopher Ibenegbu, Dr Shyam Masakapalli, Micaela Chacon, Emanuele Kendrick, Dr Leann Bacon, Ali Hussein, Luke Williams, Alexandria Holland and Maria Vittoria Ortenzi. You were so generous with sharing your experience, knowledge and kindness. Special "thank you" goes to Dr Charlotte Hamley-Bennett for talks in the early mornings, tutorials in enzyme assays and last minute help with writing this thesis but first and foremost your friendship. Thank you Lisa Buddrus for listening and being there for me. I couldn't have done this without you two.

Finally, I would like to thank my parents and my partner Charles Fane-Gladwin. This thesis is dedicated to you. Charles, you were in my corner and supported me throughout this PhD journey. Thank you my darling for being so patient, understanding and encouraging. Thank you for making me laugh when I was about to cry. I am so lucky to share my life with you.

Kochani Rodzice,
Dziękuję za Wasza miłość, wsparcie i wiarę w mój potencjał. To Wy dajecie mi siłę i pewność siebie do chwytania nowych możliwości każdego dnia. Bardzo Was kocham i ta praca jest dedykowana Wam.

Contents

List of Figures

List of Tables

Abbreviations

- `ACT`: Artemis Comparison Tool
- `BLAST`: Basic Local Alignment Search Tool
- `bp`: base pair
- `CDS`: Coding DNA Sequence
- `cobI`: precorrin-2 methyltransferase
- `cobM`: precorrin-4 methyltransferase
- `copG`: plasmid copy number regulator
- `COBRA` :The COnstraints Based Reconstruction and Analysis
- `CRISPR`: Clustered regularly-interspaced short palindromic repeats
- `dITP`: deoxyinosine triphosphate
- `DOC`: Death On Curing protein
- `EC number`: Enzyme Commission number
- `ED pathway`: EntnerDoudoroff pathway
- `FBA`: Flux Balance Analysis
- `GAM`: growth-associated maintenance
- `GAPDH`: Glyceraldehyde 3-phosphate dehydrogenase
- `GEM`: Genome-scale metabolic model
- `GO`: Gene Ontology
- `GPR`: Gene-Protein Relationship

- `HemY` protoporphyrinogen IX oxidase
- `HUS`: Hemicellulose Utilisation System
- `IPTG`: β-D-thiogalactosepyranoside
- `ITP`:inosine triphosphate
- `kb`: kilobase
- `KEGG`: Kyoto Encyclopedia of Genes and Genomes
- `LCM`: Lignocellulosic material
- `MGSA`: Model-based Gene Set Analysis
- `MOMA`: Minimisation of Metabolic Adjustment
- `nadB` L-aspartate oxidase
- `NGAM`: non-growth-associated maintenance
- `NTP`:nucleoside triphosphate
- `OD`: Optical Density
- `ODE`: Ordinary differential equation
- `ORF`: Open Reading Frame
- `Pdu`: propanediol utilisation operon
- `RAST`: Rapid Annotations using Subsystems Technology
- `repB`: plasmid replication protein
- `SDS-PAGE`: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
- `SpP`: Signal peptide type I
- `TCA` tricarboxylic acid cycle
- `XMP`: xanthosine monophosphate

Chapter 1

INTRODUCTION

1.1 Metabolic Models

Over the past decade genome-scale metabolic models have become of major interest in the scientific community. Construction of the models and their analysis have become a next logical step after sequencing of the organism of interest.Why make metabolic models of microorganisms of interest? It is common to find that the physiological responses of a microorganism in laboratory conditions is hard to explain based only on singular metabolic pathways. It is common to find that the answer to the question at hand is a combination of a variety of factors which are a result of a complex network of reactions. GEMs are a compendium of knowledge for a given organism with respect to genome annotation, biochemical properties and metabolic capabilities. Flux balance analysis allows for a gaze into metabolic networks with respect to how the reactions are distributed and which routes are the most preferential. Metabolic models allow for an *in silico* prediction of growth on a range of substrates, analysis of network perturbations in different conditions and a look into the metabolism of knock-out strains amongst others. James E. Bailey [12] explains that metabolic models aid in three main ways:

1. Uncover new and creative strategies and design
2. Think outside the conventional knowledge
3. Understand the interconnection between all the components in the system ...

However, the most important aspect is the way of looking at a metabolism in a holistic approach where perturbations to one reaction affect the whole system. From that point of view,

metabolic models are a very useful tool in an arsenal of scientific tools. Metabolic models shed light on genotype-phenotype relationships and explain bacterial behaviour observed in laboratory conditions. A model's complexity evolves with the amount of information that can be incorporated into them and as such evolves when enriched with high-throughput data [48].

In simple terms, metabolic models are biochemical networks of reactions translated into a mathematical model in the form of a stoichiometric matrix. The matrix incorporates the information on the reactions and metabolite concentrations using simple mass balance equations. Models are a compendium of knowledge available for a given organism that can be represented in a mathematical form and it is exactly that feature that makes looking into physiological properties of a system under a given environmental condition possible (using as simple a tool as flux balance analysis).

1.1.1 Spectrum of metabolic models

Describing a metabolic network of reactions calls for defining metabolite concentrations and reactions rates. Ideally, the metabolic model would include kinetic parameters however, these are difficult to determine on a genome-scale. Instead the Flux Balance Analysis is used to examine the possible metabolic network for a given species. The conservation of matter states that a closed system must have the same mass of reactants as the mass of products for the matter in a closed system must remain constant and cannot be destroyed or created. In other words, the net rate for producing a metabolite takes into calculations both the producing and consuming reactions and therefore remain in balance[47, 96, 48]. Metabolic models usually encompass hundreds of reactions which makes it difficult to have detailed kinetics data for each reaction. Not to mention that the enzyme activity and respectively reaction kinetics vary depending on physiological conditions, which makes it a challenging issue in the description of a global metabolism. To overcome these problem genome scale metabolic models are an *in silico* representation of a system in steady-state conditions where reaction rates are averaged for all rather than estimated for each enzymatic reaction. In the estimation of average rates, which are usually calculated per minute, it is not crucial to take into account the changes in regulation of the network. The network is in a steady-state with regards to the concentration of metabolites inside the cell due to the fact that those metabolites reach a high turnover rate and, if the environmental conditions stay the same, the metabolites stay at a steady level once they reach it [47, 119, 123].

1.1.2 Metabolic models constraints

Metabolic models are knowledge databases arranged in such a way that they can be easily converted into mathematical form. The system is subjected to both rules of genetics and constraints such as thermodynamics, mass and charge conservation and substrate / enzyme availability [72, 96].

1. The enzyme and substrate availability is in the context of genomic-scale metabolic models understood in terms of single reactions: it is the availability of substrate either from defined cell environments or produced by other reactions and the presence of corresponding enzymes to facilitate a given reaction.

2. Mass and charge conservation relates to a stoichiometric matrix which is the mathematical heart of the model and limits the number of substrates and products for each reaction in the matrix.

3. The thermodynamic rules are responsible for the directionality of each reaction.

4. Finally, the reaction flux bounds can constrained using experimental data.

These rules main role is to constrain the solution space to describe the most appropriate representation of the the flux distribution. It is a multidimensional representation of the fluxes present in the whole metabolic reconstruction in a steady-state (see Figure 1.1).

Those rules when satisfied paint a comprehensive genotype-phenotype relationship specific to the organism of interest.

1.1.3 Mathematical approach to metabolic modelling

The complex network of reactions in a given species is represented in a form of a stoichiometric matrix (S) where each reaction can be found in columns (Rxn1-8) and corresponding reactants and products in rows (A-I). Stoichiometric coefficients are mathematical equivalents of input and output for each reaction [72, 119, 96].

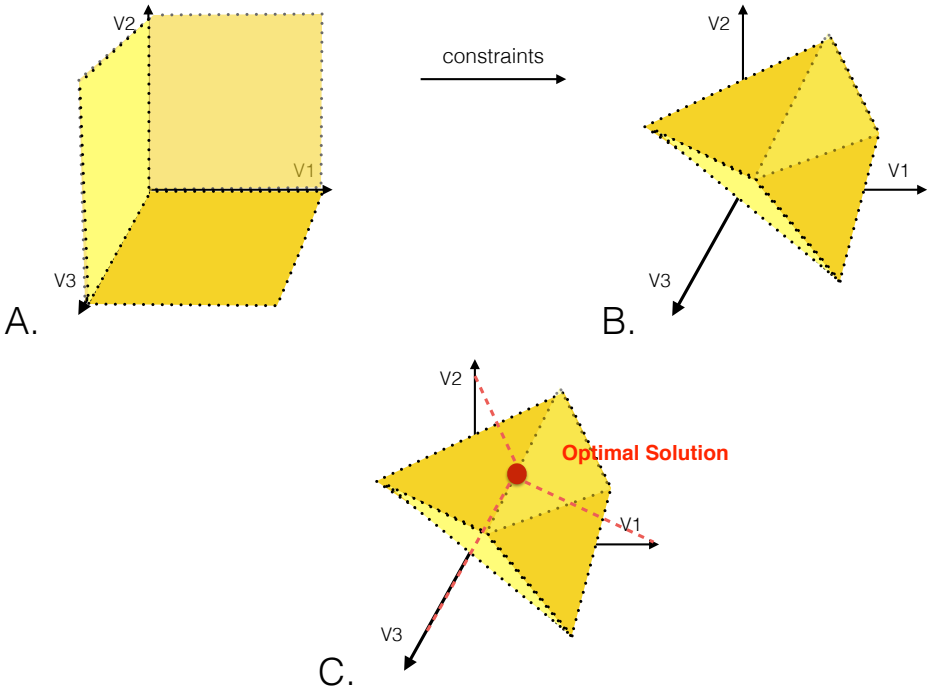


Figure 1.1: The visual representation of solution space and the stages of constraining the model. A. represents unconstrained solution space, B. shows constrained metabolic solution space and C. shows the optimum solution for the model.

$$
S = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \\ G \\ H \\ I \end{array}
\begin{array}{c}
\begin{array}{cccccccc} Rxn1 & Rxn2 & Rxn3 & Rxn4 & Rxn5 & Rxn6 & Rxn7 & Rxn8 \end{array} \\
\left(\begin{array}{cccccccc}
-1 & 0 & 0 & 1 & 0 & -1 & 0 & -1 \\
1 & -1 & 0 & 0 & 1 & -1 & 1 & 0 \\
0 & 1 & -1 & 0 & -1 & 0 & 0 & 1 \\
1 & 0 & -1 & -1 & 0 & 1 & -1 & -1 \\
0 & 1 & -1 & 2 & 1 & 1 & -1 & 0 \\
0 & 1 & 0 & -1 & 1 & 1 & 0 & -1 \\
-2 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \\
1 & -2 & 1 & -1 & 0 & 0 & -1 & 0 \\
0 & 0 & 1 & 2 & 0 & -2 & 1 & 1
\end{array}\right)
\end{array}
$$

Each reaction is described by stoichiometric coefficients representing a number of substrate and product molecules participating in the reaction. However, stoichiometric coefficients can change from negative to positive value depending on the directionality of the reaction [67, 124]. This can be demonstrated by the following example. If we assume that the reaction

is:

$$1S_1 + 3S_2 \rightarrow 2P_1 + 2P_2, \tag{1.1}$$

,where S_n denotes a unique substrate and P_n product, then the corresponding coefficients for this reaction would be: -1, -3, 2, 2. However, if the reaction would be happening in an opposite directions the coefficients would be: 1, 3, -2, -2.

Each intracellular metabolite that is found in the stoichiometric matrix can be described using ordinary differential equations (ODE) where the rate of degradation and subsequently production of a given metabolite is taken into account (Eq. 1.2) [67, 23].

$$\frac{dx}{dt} = v \tag{1.2}$$

This equation would change depending on stoichiometric coefficient, for instance, for Eq. (1.1) the Eq. (1.2) would change to:

$$\frac{dS_1}{dt} = -v, \quad \frac{dS_2}{dt} = -3v, \quad \frac{dP_1}{dt} = 2v, \quad \frac{dP_2}{dt} = 2v \tag{1.3}$$

which also suggests that the rate of degradation of S_2 is higher than that of S_1 and the rate of production for both products P_1 and P_2 is the same.

Metabolic models are mathematically described in a form of a system or balance equations which means the above equation for a single reaction can also describe the entire metabolic network Eq.(1.4)[67].

$$\frac{dXi}{dt} = \sum_{j=1} n_{\mathrm{ij}} v_{\mathrm{j}} \quad for \quad i = 1, ..., m \quad and \quad j = 1, ..., r. \tag{1.4}$$

Here j refers to the reaction number and i to a corresponding metabolite. The n_{ij} denotes the stoichiometric coefficients for a relevant reaction [67]. X refers to metabolites (either substrate or product).

For mathematical representation of metabolic model hence we need to define the stoichiometric matrix `N` , `v` as a vector of reaction rates, `X` as the vector of concentrations and a vector of parameters `p` along with vector`J` containing fluxes [67]. The resulting equation derived from the balance equation would be:

$$\frac{dX}{dt} = Nv \tag{1.5}$$

and for the steady state assumption the equation would be:

$$\frac{dX}{dt} = Nv = 0 \tag{1.6}$$

Figure 1.2 shows an example of a network of three reactions with complete set of corresponding ODEs, where the above equations and steady state assumptions are used in practice.

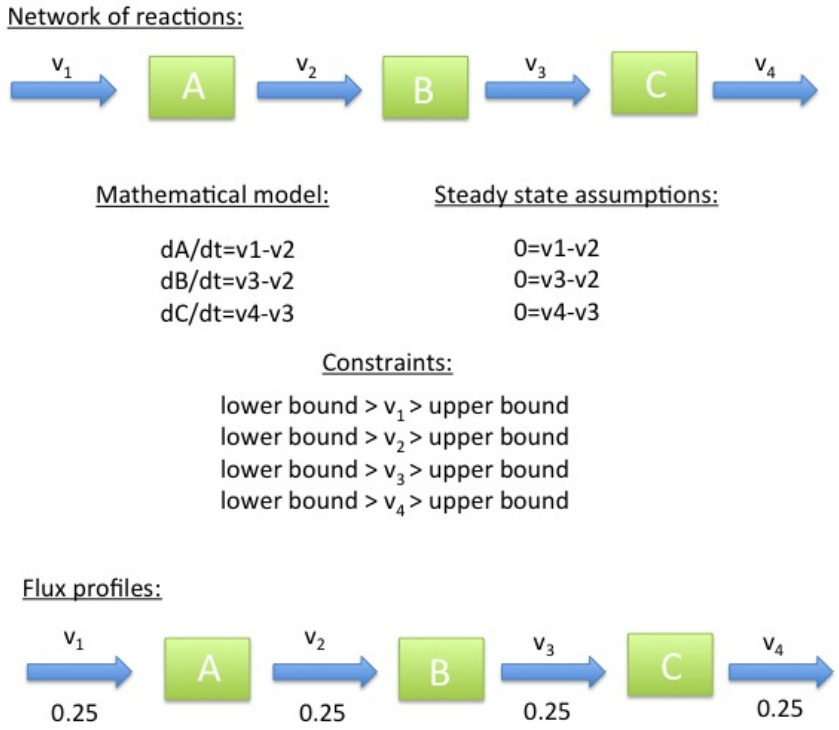


Figure 1.2: An example of mathematical representation of a hypothetical network of reactions with corresponding ODEs, constraints and steady state assumptions.

This mathematical representation allows for calculation of possible fluxes within the network and detection of dead ends and unbranched pathways alike [67]. It should be noted that in the assumption of a steady state, all the rates of production and degradation of a given metabolite must have a net value of 0 [67].

Each reaction, apart from satisfying mass balance requirements, can be further constrained with regards to upper and lower bounds. These bounds reflect the directionality and reversibility of a given reaction and can be deduced based on thermodynamic laws where an

irreversible reaction has a non-negative reaction rate [47]. The irreversibility of a given reaction can be either known from *in vivo* observations or deduced using laws of thermodynamics, such as Gibbs free energy. [96, 47].

The resulting flux is represented in units: mol per hour per gram of dry weight. Flux distribution is defined in metabolic modelling as the network of reactions and collective flux of a system that they represent [47, 96].

The system is defined as a whole cell with the surrounding environment and incorporated into a mathematical framework. The system can uptake some metabolites from the environment via transporters and only some can be excreted, as a collective those metabolites are defined as external metabolites (see Figure1.3). Intracellular metabolites are defined as the metabolites used in the network of reactions and are included in the stoichiometric matrix whereas, external metabolites are not.

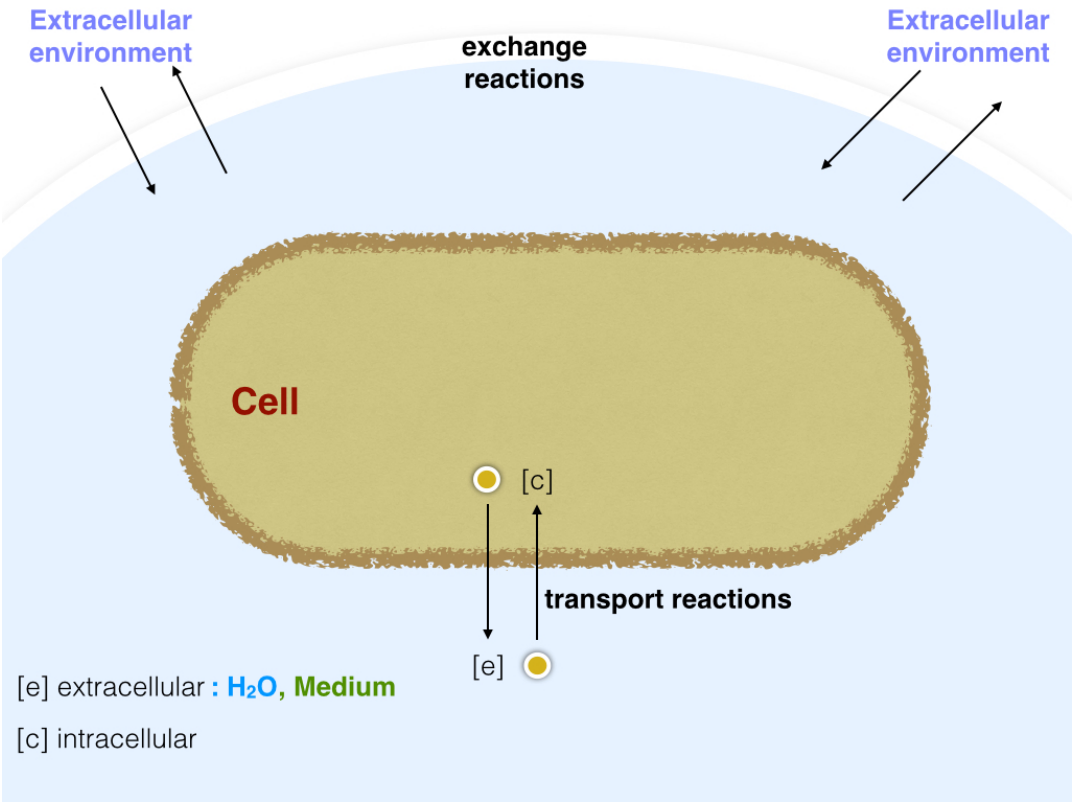


Figure 1.3: Representation of the extra- and intracellular metabolites in genome scale metabolic models with relation to transport and exchange reactions.

Metabolic models are considered correctly assembled when all the above criteria are met which is mirrored by feasibility of fluxes which has a direct correlation with a previously described linear equation [47, 96]. Construction of genome-scale metabolic models to highest standards and quality requires an iterative approach with incorporation of a vast amount of

relevant datasets [134, 96]. The approach of incorporating biochemical, genetic and genomic knowledge gives a unique level of understanding for a given organism, which is why creating metabolic models have become a natural step after genome sequencing of a target species [134].

1.2 *Geobacillus* spp

Geobacillus spp. can be found in a variety of environments ranging from a thermal bore-hole pipe in the Southern Urals through oilfields in China to rotting wood in Florida, USA [118]. Strains now classified as members of genus *Geobacillus* have initially been classified by Ash *et al*, based on 16S rRNA as belonging to "group 5" in the taxon *Bacillus*. Thanks to work of N. A. Nazina at the Russian Academy of Science, who was interested in microbial communities found within high temperature oilfields in Kazakhstan, a new genus *Geobacillus* was introduced based on painstaking examination of the isolates found there. The new group was officially introduced in 2001 and was described as rod-shaped aerobic or facultative anaerobic thermophilic spore-forming bacilli (oxygen can be replaced as an electron acceptor by nitrate in strains such as *G. thermodenitrificans*).

Nazina *et al* [136] described the new taxa with regards to morphology, metabolism, DNA characterisation and ecology. *Geobacillus* spp. grow in temperature range from 37-75℃(optima at: 55-65℃) with pH 6.0-8.5 [136], although they can also be found in environments with temperatures as moderate as 30℃[60]. This is accredited to the resilience of spores and their ability to travel and remain suspended in air due to the size smaller than 1 µm [7]. Indeed, spores found in taxa *Bacillus* and *Geobacillus* display a high resilience to UV light, chemical or thermal treatment and as such have been used in analysing effectiveness of antiseptic techniques [139]. The successful widespread of *Geobacillus* spp. is a testimony to a genus which grows vigorously in composting plant material, thriving in high temperature environments and owes its survival to small, resilient spores which can remain airborne for extensive periods of time and if dormant remain viable [60]. This explains the abundance of *Geobacillus* spp. in temperate climates, not only on the surface but also deep in the soil layers [111].

Sequencing efforts which rendered over 30 *Geobacillus* genome sequences made it possible to explore the origins of thermophily in this genus however these efforts are still ongoing. Works by Takami *et al*, 2007 or Feng *et al*, 2002 were able to narrow down a few features which were uniquely present in the thermophilic bacilli [144, 51]. *Geobacilli* ensure RNA and DNA stability in high temperatures through use of protamines, spermidines, spermines, tRNA methyltransferase and spermidine synthase [144]. To date however little has been found about the origin of the thermophilic features in *Geobacillus* spp. What has been observed within the genus is the large number of transposable elements present within the species

across the clades, which might suggest transfer of thermophilic genes from other strains. BLAST analysis on a few thermophilic genes suggests however a closer evolutionary route from taxa such as *Anomybacillus* or *Paenibacillus* rather than *Archaea* as it was previously suggested in the literature [111].

The other features which were picked up by other reviews of this genus, were the high number of CRISPR elements present within the genomes of these thermophilic bacilli. CRISPR and Cas-proteins associated with them are described as the immune system of microbes [111, 60]. Even though a high number of those elements can be present within the genomes of thermophilic bacilli, compared to their mesophilic counterparts, the actual mechanism in which they would assist in thermophily remains unclear. Those elements might however suggest that the genus is still undergoing active evolution, which in turn might explain the differences between the clades within the genus or indeed on a sub-strain level.

Due to their biotechnological potential and peculiarity of the metabolism, the group has received significant attention from academia and industry alike.

Infamously, *Geobacillus* spp. have been culprits of dairy contamination [148].

1.2.1 Biotechnological application of *Geobacillus* spp.

Geobacillus spp. can be used both in whole-cell applications and in biofuel and chemical production through engineered cells. The former is applied in extraction of long linen fibres which in turn can be used in textiles, insulating materials or papers. One of the main advantages of using bacteria from this taxa is faster rate of reaction, decreased contamination and easier maintenance [127, 38]. *Geobacillus* spp. can be used as factories for multiple products; from gold nanoparticles using *Geobacillus* sp. strain ID17 (using NADH-dependent enzymes which convert Au3+ to elemental gold) [37] to production of bioethanol by *Geobacillus thermoglucosidasius* TM242. The members of the genus have a strong application in bioremediation, especially with regards to degradation of aromatic compounds or organophosphates (the latter activity was found in *G. caldoxylosilyticus* T20), as well as in biocontrol with relation to biocontrol of *Fusarium* wilt [37, 60].

Most prominently however *Geobacillus* spp. have been used in strain metabolic engineering for production of biofuels and biochemicals [60, 38]. *Geobacillus* spp. have been a well of thermostable enzymes used in the industry, such as lipase, glycoside hydrolases, N-acylhomoserine lactones, DNA polymerase or protease to name but a few. Hydrocarbon degradation is one of the widespread features of the genus *Geobacillus*. This is not surprising, considering the

fact that a few of the species were found growing in oilfields. There is also no clear preference towards the chain-length of alkanes degraded or the type of substrates [82].

G. thermoleovorans is able to degrade compounds ranging from pentane to nonadecane, along with polycyclic aromatic hydrocarbon naphthalene. Indeed, members of this genus are able to efficiently grow on as diverse substrates as: long-chain alkanes, cellulose, hemicellulose or polyvinyl alcohol [127]. They can withstand arsenic and high levels of ethanol and herbicides [127, 98]. The clear advantages include: a simple way of removing unwanted products, maintaining aerobiosis with added benefit of keeping ethanol concentration below lethal levels for the bacterium not to mention reduction in the cost of cooling systems and decreased possibility of contamination with pathogens. In short they are perfectly equipped to perform high-temperature microbial processes in sewage or waste treatment [127]. It is no wonder that a few *Geobacillus* spp. have been applied in high-temperature microbial processes including production of acetone and 2,3-butanediol using *Geobacillus* sp. XT15 or bioethanol production by *Geobacillus thermoglucosidasius* [127].

1.2.2 *Geobacillus thermoglucosidasius* NCIMB 11955

The motivation for this research was to investigate the metabolism of *Geobacillus thermoglucosidasius* NCIMB 11955 which as a thermophilic bacterium has tremendous importance as a platform for production of bioethanol. Thermophily comes as an advantage in the process of purification of bioethanol where high temperature plays a crucial role. Furthermore, *Geobacillus thermoglucosidasius* can use carbohydrate components of lignocellulose for growth which allows for utilization of lignocellulose rather than nutritiously important feedstock [127, 60]. On an industrial scale for production of highly desirable products factors such as: yield, titer, robustness and productivity play a major role and *G. thermoglucosidasius* NCIMB 11955 has been able to satisfy those criteria [127, 60]. This is precisely a reason why, paired with metabolic capabilities, this bacterium has been used already on an industrial scale by companies such as TMORenewables Ltd [38].

1.2.3 Motivation

The motivation of this work was to explore the metabolic capabilities of *Geobacillus thermoglucosidasius* NCIMB 11955. This work aims to provide, through metabolic model and output from tools such as PathwayBooster, a comprehensive compendium of up-to-date metabolic data available for this bacterium. Although this work is heavily reliant on bioinformatics tools, it also benefits from the data generated through experimental work by the author of this thesis, Dr Shyam Masakapalli, Dr Leann Bacon, Dr Steven Bowden, Carolyn Williamson, and Dr Alice Marriott. The sequenced genome of *Geobacillus thermoglucosidasius* NCIMB 11955 was the starting point for the reconstruction of the genome-scale metabolic model and was done by ERGO Integrated Genomics [104]. Although this bacterium has been studied extensively by a research group under Professor David Leak at University of Bath, and central carbon metabolism C13-base flux analysis [130] were created for *Geobacillus thermoglucosidasius* the objective of creating a genome-scale metabolic model for this strain was truly novel. The genus *Geobacillus* and its metabolic capabilities were studied on strains such as *G. stearothermophilus* [88], *G. kaustophilus* [126], *G. thermodenitrificans* [31]. A comparative analyses on various strains within this genus were published on the hemicellulose utilization loci (HUS) [45] and the origins of thermophily [126] however an extensive comparison between the metabolic capabilities of these strains and *G. thermoglucosidasius* was not. The objective to this comprehensive analysis was to understand the underlying differences between the strains that are observed in the laboratory and understand the limitations of *G. thermoglucosidasius* NCIMB 11955 such as dwarfed growth in strictly anaerobic conditions as well as explore the potential *in silico* applications of this bacterium in production of succinate or butan-2,3-diol.

Chapter 2

G. thermoglucosidasius NCIMB 11955: structural features.

2.1 Introduction

The motivation for this research project was to understand *Geobacillus thermoglucosidasius* NCIMB 11955 in more depth both in its metabolic capabilities and limitations. The genus *Geobacillus* can be divided roughly into five clades (see Figure 2.2); *Geobacillus thermoglucosidasius* NCIMB 11955 belongs to clade "thermoglucosidasius". At the beginning of this research project the only other sequenced genome in this clade was that of *G. thermoglucosidasius* C56-YS93, however the corresponding paper on its metabolic capabilities was not published at the time. This study began with characterisation of genome organisation and gene arrangement of the strain and its comparison to other *Geobacillus* strains both within the clade "thermoglucosidasius" and the remaining four clades. In this chapter, the genome rearrangements between strains across the clades and within the clade "thermoglucosidasius" is analysed. During this part of the research, it was found that there are three highly conserved regions shared across the genus. These conserved regions were found to code, amongst others, for hemicellulose degradation, vitamin B-12 metabolism and enzymes involved in central carbon metabolism.The study of hemicellulose degradation locus (HUS) was previous described in publication by DeMaayer, 2014 [45], however it lacks the analysis of HUS locus in clade "thermoglucosidasius" that this research project includes.

This global study provides a novel insight into the clade "thermoglucosidasius" and compar-

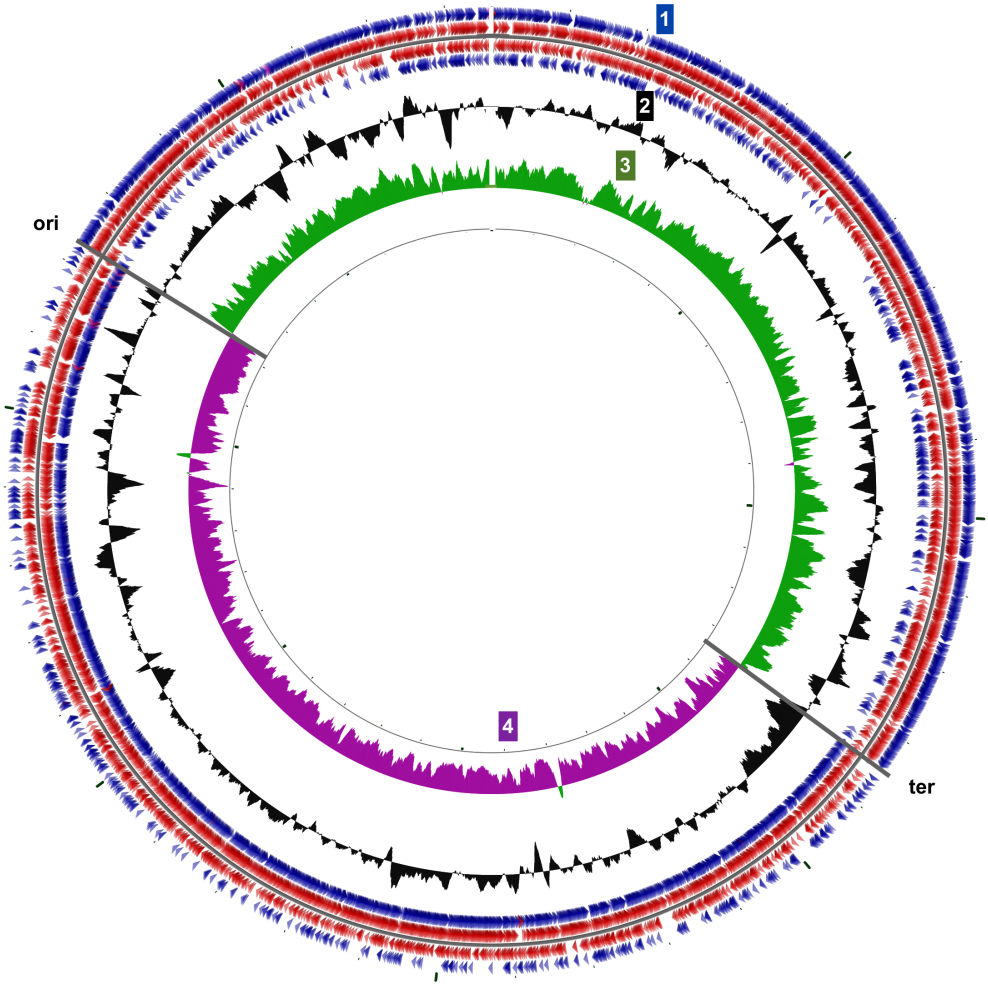


Figure 2.1: Circular view of the chromosome of *G. thermoglucosidasius* NCIMB 11955. [1] denotes : CDS [2]: GC content,[3] : positive GC skew and [4]: negative GC skew.

ison between *G. thermoglucosidasius* NCIMB 11955 and other strains across the genus. The analyses of the genomic regions of high conservation shows the core metabolic capabilities of the genus, with respect to adaptation to thermophily and underlying core metabolism. The focus on understanding how the genomic diversity occurred led to the analysis of mobile genetic elements, which is why in this chapter genome characteristics such as CRISPR regions, transposons and presence of phages are discussed.

2.1.1 Clades within the genus *Geobacillus*

Originally the classification of thermophilic *Geobacillus* spp. as a distinct genus from other *Bacillus* spp. was based on analysis of 16S sequences. Ziegler *et al* [150] have argued that an analysis based on the sequence of the *recN* gene, figure 2.2, shows greater phylogenetic discrimination between strains present in taxa *Geobacillus*. This approach provides a higher degree of distinction between individual species. The *Geobacillus* spp. can first and foremost be divided into facultative anaerobes and strict aerobes. This distinction is mirrored by an early division in the phylogenetic tree (Figure 2.2). The taxa *Geobacillus* then diverges into five clades; namely: "kaustophilus", "stearothermophius", "thermodenitrificans", "caldoxylosilyticus" and "thermoglucosidasius". Clade "kaustophilus" encompasses a number of proposed species including *G. thermocatenulatus*, *G. caldotenax*, *G. litanicus* or *G. vulcani*. To elucidate subtle distinctions between proposed strains within the clade a single-nucleotide variation analysis should be used on a group of conserved genes. Such an approach would be beneficial especially given the number of sequenced genomes now available from this clade.

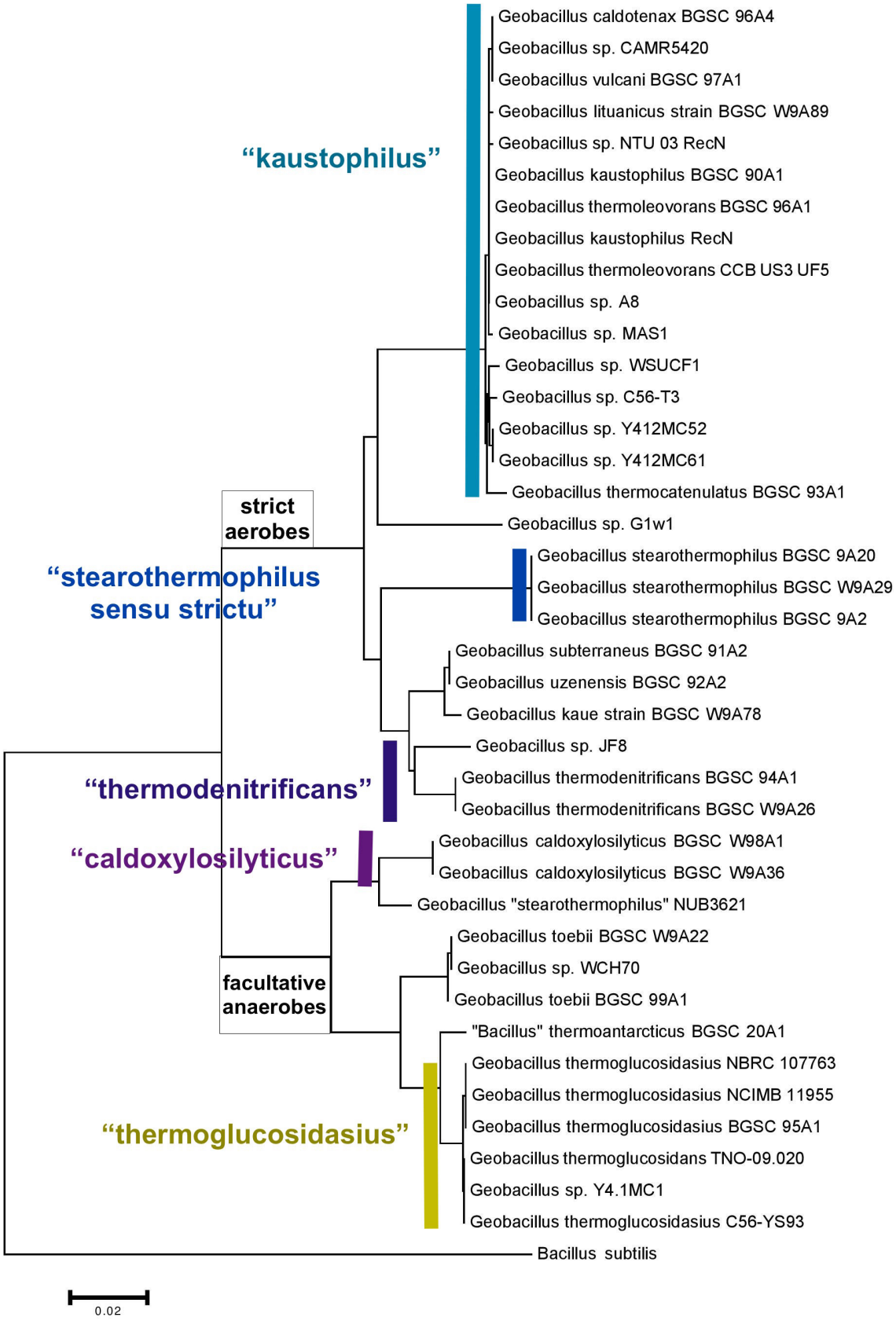


Figure 2.2: Evolutionary relationships of taxa *Geobacillus* based on *recN* gene. The evolutionary history was inferred using the Neighbor-Joining method [108]. The evolutionary distances were computed using the Maximum Composite Likelihood method [5]. Evolutionary analyses were conducted in MEGA6 [129].

2.2 Genome Organisation

2.2.1 Plasmids

The genome sequence of *G. thermoglucosidasius* NCIMB 11955 clearly shows the presence of two mega plasmids: pGTH11955-01 (83,925 bp) and pGTH11955-02 (47,897 bp). It is unusual to find two mega plasmids in strains of *Geobacillus*. Only a few other strains have been found to have plasmids and even fewer have been characterised. Plasmids have been found in: *G. kaustophilus* HTA426 (pHTA426), *G. stearothermophilus* STK (pSTK1), *G. thermodenitrificans* NG8O (pNG8O-2),*Geobacillus* sp. Y4.1MC1 (pGY4MC101), *Geobacillus* sp. 610 (pGTD7), *Geobacillus* sp. 1121 (pGTG5) and *G. thermoglucosidasius* C56-YS93 (pGEOTH01, pGEOTH02) [60]. It should be noted that only substrains of *G. thermoglucosidasius* has been found to encode two mega plasmids, however those have not been characterised in the literature. A detailed analysis of the genes encoded on the plasmids of *G. thermoglucosidasius* NCIMB 11955, was therefore potentially important to understand the advantages these two mega plasmids bring to the strain. A search for the plasmid replication protein has shown the presence of a *repB* gene. RepB has been characterised as the initiator protein for rolling-circle type of replication [76]. Transcription of *repB* has been shown to be under strict control of translation initiation regulatory signals and two trans-acting elements, namely CopG and the antisense RNAII [76]. Upstream region from the *repB* gene is highly conserved sequence, which has been shown to play a major role in the correct translation, whereas translation is dependent on *repB* own initiation signal rather than being coupled with that of *copG* [76]. Moreover the sequence upstream from *repB* coupled with the atypical ribosomal binding site proximal box along with four bases directly downstream of it, plays a crucial role in the translation of *repB* [76]. This sequence has also been found to be conserved amongst the *repB* region found on other *Geobacillus* plasmids. BLAST search has suggested highest sequence similarity of rep region between that of pSTK1 from *G. stearothermophilus* [88] and other plasmids found within the strains of genus *Geobacillus*, which has not been characterised before.

Both plasmids were found to encode a large number of transposases, integrases and IS elements, which in themselves might explain difference of size and gene content between the plasmids of *G. thermoglucosidasius* NCIMB 11955 and C56-YS93 and indeed between the plasmids found in the genus *Geobacillus*. A complete list of elements found on the plas-

mids can be viewed in the appendices however a few genes are worth highlighting, such as a gene encoding Death ON curing protein (DOC) and the associated Phd encoding gene on pGTH11955-01. These proteins are responsible for maintaining stable plasmid inheritance by a post-segregational killing system [71, 140]. The plasmids also harbour genes which appear to encode bacitracin and chloramphenicol resistance, but given that *G. thermoglucosidasius* TM242 is sensitive to the latter, these would need to be artificially induced to confirm the activity.

Table 2.1 shows a selection of plasmid-borne genes which provide unique metabolic capabilities to *G. thermoglucosidasius* NCIMB 11955 and are not found in other species in this genus, characterised to date. It is however clear that the genes found on the plasmid bring benefit to the strain in adverse environments and broadens the spectrum of substrates it can utilise as carbon source, including catechol degradation (catechol-2,3-dioxygenase, 4-oxalocrotonate tautomerase) which will be discuss in more detail in the following chapter.

EC number	Description	Prevalence in the genus
EC 1.1.1.95	D-3-phosphoglycerate dehydrogenase	*G. thermoglucosidasius* C56-YS93
EC 1.13.11.2	Catechol-2,3-dioxygenase	*G. thermoglucosidasius* C56-YS93
EC 1.2.1.10	Acetaldehyde dehydrogenase	absent in *G. kaustophilus*
EC 2.3.1.37	5-aminolevulinic acid synthase	*G. thermoglucosidasius* C56-YS93
EC 4.2.1.80	2-oxopent-4-enoate hydratase	*G. thermodenitrificans* CCB-US.3-UF5
EC 4.1.3.39	4-hydroxy-2-oxovalerate aldolase	*G. thermodenitrificans* CCB-US.3-UF5
EC 5.3.2.6	4-oxalocrotonate tautomerase	*G. thermoglucosidasius* C56-YS93
EC 4.1.1.77	4-oxalocrotonate decarboxylase	*G. thermodenitrificans* CCB-US.3-UF5
EC 2.4.21.89	Signal peptidase I	from *Bacillus subtilis* plasmid

Table 2.1: Table showing selection of interesting genes found on two mega-plasmids of *G. thermoglucosidasius* 11955: pGTH11955-01 and pGTH11955-02.

The signal peptidase type I (Spases SpP) gene present on pGTH11955-01 is also found on a *Bacillus subtilis* plasmid. This plasmid-borne peptidase was shown to replace the main chromosomally encoded SipS and SipP and reaches its maximum expression level during the post-exponential growth phase [135].

A 5-aminolevulinic acid synthase gene appears to be found uniquely in the strains of *G. thermoglucosidasius*, which suggests that in these strains, synthesis of 5-aminolevulinic acid is from glycine rather than L-glutamate. This precursor plays a bottleneck role in the synthesis of vitamin $B_{12}$and protoheme [107].

Genes encoding 2-oxopent-4-enoate hydratase along with 4-hydroxy-2-oxovalerate aldolase and 4-oxalocrotonate decarboxylase are found in a gene cluster on the plasmid, are homologous to genes found on *G. stearothermophilus* strain DSMZ 6285 plasmid pGGO1 and in the chromosome of *G. thermodenitrificans* CCB-US.3-UF5 (see Figure 2.2). These en-

EC number	*G. stearothermophilus* DSMZ 6285	*G. thermodenitrificans* CCB-US.3-UF5
EC 4.2.1.80	99%	69.60%
EC 4.1.3.39	99%	81.55 %
EC 4.1.1.77	99%	59.30 %

Table 2.2: Table showing the sequence similarityT results for 2-oxopent-4-enoate hydratase **EC 4.2.1.80**, 4-hydroxy-2-oxovalerate aldolase **EC 4.1.3.39** and 4-oxalocrotonate decarboxylase **EC 4.1.1.77** to *G. stearothermophilus* strain DSMZ 6285 and *G. thermodenitrificans* CCB-US.3-UF5. The sequence similarity was calculated through BLAST.

zymes play a crucial role in catechol degradation allowing the metabolic conversion of 2-oxopent-4-enoate through 4-hydroxy-2-oxopentanoate to acetaldehyde and pyruvate. This gene cluster provides a complementary pathway to the activity of 2-hydroxymuconic semialdehyde hydrolase (EC 3.7.1.9) encoded on the chromosome of *G. thermoglucosidasius* NCIMB 11955 which facilitates the second step in the breakdown of catechol, converting of 2-hydroxymuconatesemialdehyde to 2-oxopent-4-enoate. A BLAST search suggests that this hybrid chromosome-plasmid encoded catechol degradation pathway may have come from *G. stearotheromophilus* DSMZ 6285 plasmid pGGO1. Such a hybrid pathways are well described in gram-negative bacteria, but have not been extensively reported in gram-positive bacteria.

2.2.2 Genome features of *G. thermoglucosidasius* NCIMB 11955

The chromosomal G+C content of *G. thermoglucosidasius* NCIMB 11955 was compared to that of other members of genus *Geobacillus* along with genome size. The findings suggest (see Figure 2.4) that there is a correlation between genome size and G+C content within the genus. It seems that the smaller the genome, the higher the GC and conversely the biggest genomes are found to have a of relatively low G+C content.

It should be noted that this trend does not support a model of linear correlation but merely suggests a grouping of genomes with size over 3.7 Mb for propensity of G+C content around the value of 44% and similarly for genomes smaller that 3.6 Mb the GC% ranges between 52-53%. It is worth noting that the cluster of genomes with highest GC% and smallest

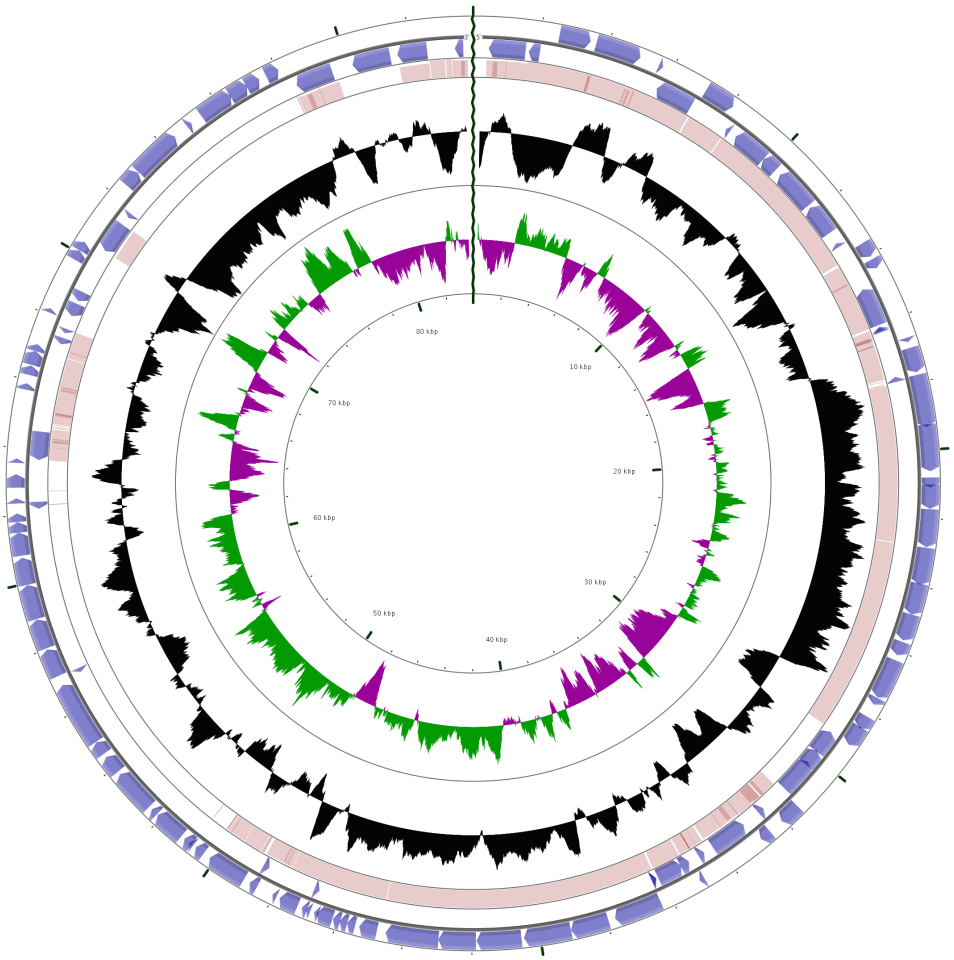


Figure 2.3: BLAST comparison between plasmids pGTH11955-01 from *G. thermoglucosidasius* NCIMB 11955 and pGEOTH-01 from C56-YS93.Purple denotes coding DNA sequence in pGTH11955-01, red: tRNA, pink: rRNA, grey: other, salmon BLAST result indicating the homologous sequences in pGEOTH-01 , black, GC content, green: GC skew + and magenta: GC skew -.

genomes all belong to clade "kaustophilus" ([126]) whereas the extremes of the opposite spectrum correspond to group "thermoglucosidasius", which *G. thermoglucosidasius* NCIMB 11955 belongs to and two genomes from species *caldoxylosilyticus*. The single data point in-between the two pronounced genome clusters belongs to *G. thermodenitrificans* NG8O-2. The difference in size within the genus *Geobacillus* has been attributed to the amount of CRISPR regions along with transposable elements ([60]). For example *G. thermoglucosidasius* strain C56-YS93 encodes 112 CRISPR-associated proteins whilst *G. kaustophilus* HTA426 encodes one CRISPR-associated helicase that is also annotated in the genome of *G. thermoleovorans* CCB_US.3_UF5 [60]. However, the most profound distinction between species that belong to groups "thermoglucosidasius" and "caldoxylosilyticus" is the adaptation to anaerobic growth

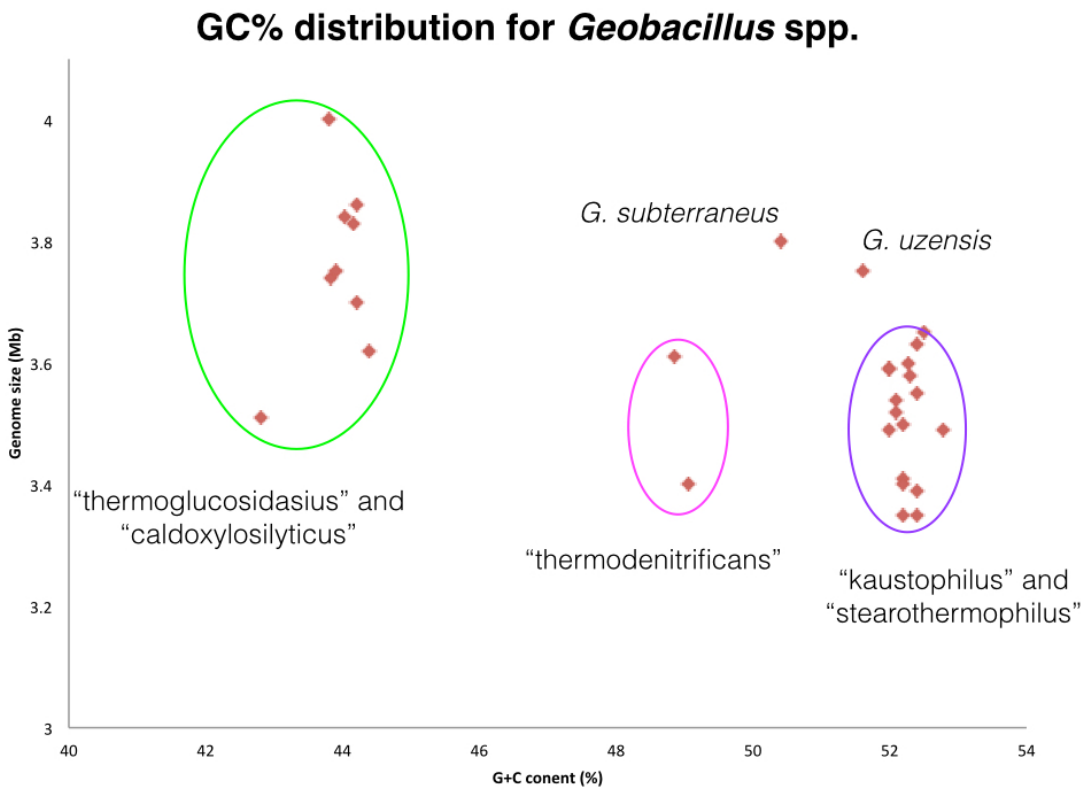


Figure 2.4: Clusters of *Geobacillus* spp strains based on their genome size and G+C content. Strains used for this analyses, along with their accession numbers can be found in " Materials and Methods".

both of which group facultative anaerobic bacteria 2.4.

G. thermoglucosidasius TM242 was analysed with regards to codon preference within the clade "thermoglucosidasius" (as defined by [126]). For the purpose of comparative analysis, codon preference was also investigated for clade "kaustophilus" (as described in chapter: Materials and Methods). Although there was no correlation between specific codon preference within clade "thermoglucosidasius" and "kaustophilus", the former group shows higher diversity in codon usage than the latter (with notable exception for clade "kaustophilus" with relation to stop codon preference). No clear preference could be observed for sub-strain comparison within the two clades. Testing a hypothesis that significant difference in codon usage might reflect an adaptation of cognate tRNA with optimal codons. The analysis was done using percentage of codon use per amino acid and frequency of codon occurrence per 1000 codons. Initially, codon usage was calculated for the strains of: *G. thermoglucosidasius* NCIMB 11955, *G. thermoglucosidasius* C56-YS93, *G. thermoglucosidans* NTO-09.020, *G. kaustophilus* HTA426 and *G. kaustophilus* Gblys, *G. thermoleovorans*CCB_US.3_UF5, *G.*

thermocatenulans BGSC 93A1 and and *G. thermodenitrificans* NG8O-2 to reflect strains present within clades "thermoglucosidasius" and "kaustophilus" (see Figure 2.5).

However, the variation of percentage of codon usage between strains led to the analysis of codon usage and subsequent calculations of standard deviation for sub-strains in the strains of *G. kaustophilus* and *G. kaustophilus*. The rationale was to see if the same variation of codon usage is observed among strain in the same clade or can this be also observed between strains evolutionarily closest to one another. Standard deviation was used as a measure of the how close the the data points deviate from the codon usage mean.

For the group "thermoglucosidasius" the codon frequency and codon percentage the strains used were *G. thermoglucosidasius* NCIMB 11955, *G. thermoglucosidasius* C56-YS93 and *G. thermoglucosidans* NTO-09.020. For the group "kaustophilus" the strains used were: *G. kaustophilus* HTA426 and *G. kaustophilus* Gblys. The term group reflects sub-strains for the strains of *G. thermoglucosidasius* and *G. kaustophilus*.

The analysis objective was to test how similar the codon usage was between the two groups of sub-strains. The higher standard deviation observed for codon within the group of "thermoglucosidasius" is an indication of wider range of codon use in this group. The possible explanation for the high standard deviation observed in the group "thermoglucosidasius" is that the codon preference might be subjected to the mutation-selection-drift theory, where in a finite population, the selection and mutation forces affect the efficiency of translation [22].

2.2.3 Genome rearrangements and comparative analyses

The genome structures of *G. kaustophilus* HTA426, *G. thermodenitrificans* NG8O-2, *G. thermoleovorans* CCB-US.3-UF5 were compared in a pairwise manner to that of *G. thermoglucosidasius* NCIMB 11955 (Figure 2.7).

This shows that there has been a significant genome rearrangement in clade "thermoglucosidasius" compared to the other species in other clades. Such a genome rearrangement might have given rise to the early divergence in the genus *Geobacillus*. When genome arrangement was analysed in a pairwise manner, comparing *G. kaustophilus* HTA426 with *G. thermoleovorans* CCB-US.3-UF5 (Figure 2.8(a)) and *G. thermoleovorans* CCB-US.3-UF5 with *G. thermodenitrificans* NG80-2 (as shown in Figure 2.8 (b)) a higher degree of genome conservation was observed. The highest genome conservation was found between the closely related *G. kaustophilus* HTA426 and *G. thermoleovorans* CCB-US.3-UF5 (Figure 2.8(a)).

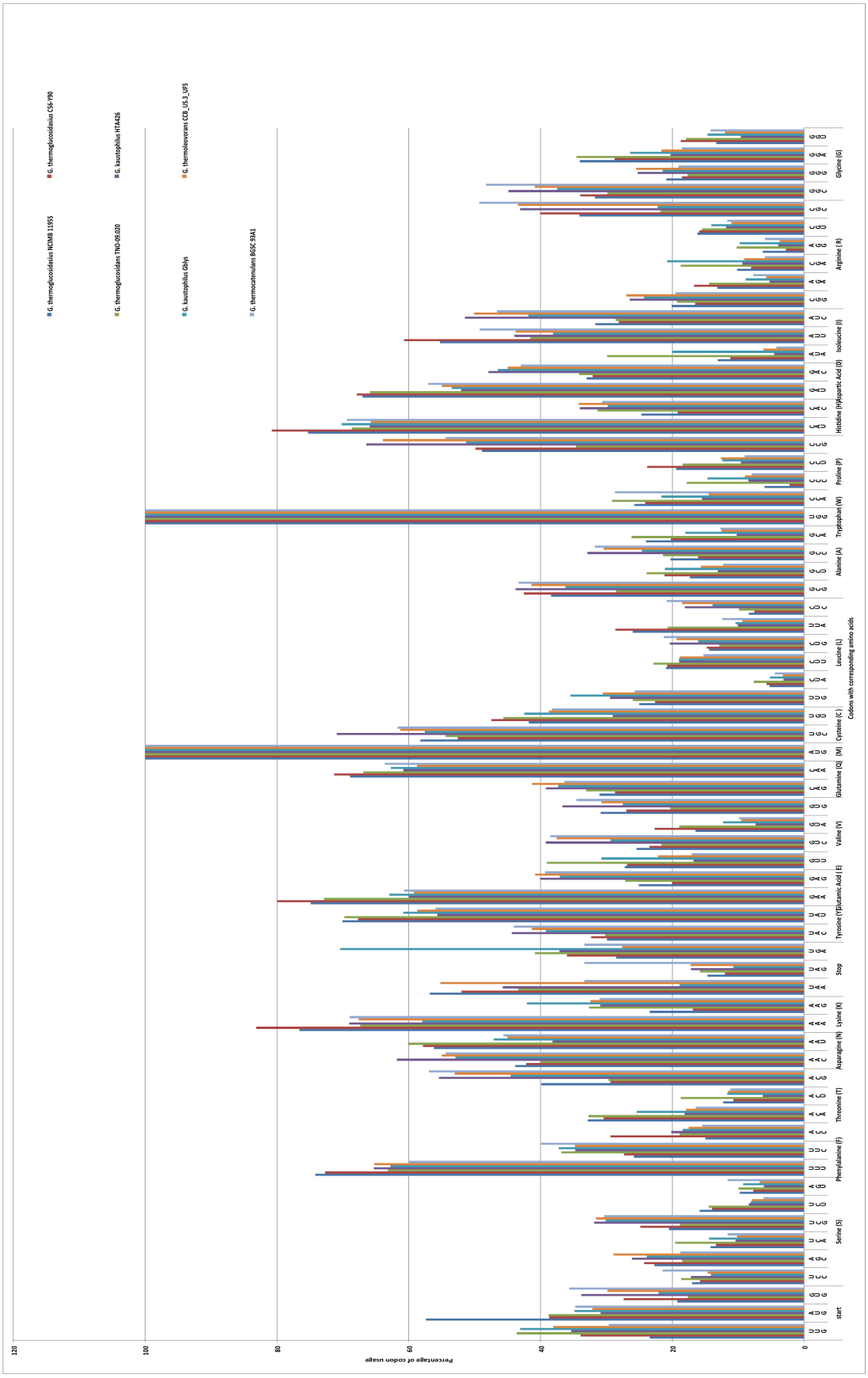


Figure 2.5: Codon usage preference for clades "thermoglucosidasius" and "kaustophilus" in genus *Geobacillus*.

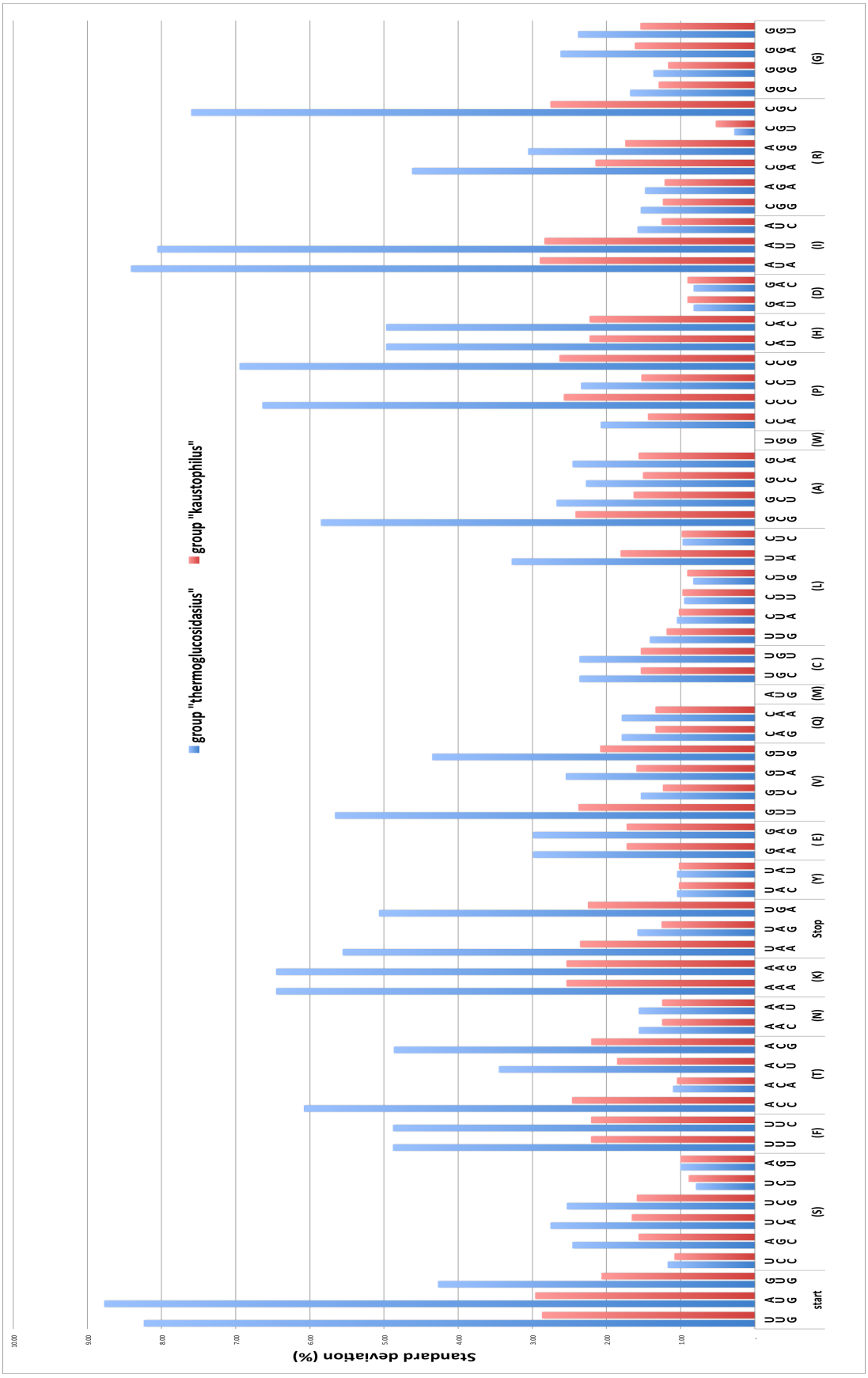


Figure 2.6: Codon usage preference (standard deviation) for group "thermoglucosidasius" and "kaustophilus" in genus *Geobacillus*. Strains used for this analyses, along with their accession numbers can be found in chapter: Materials and Methods.

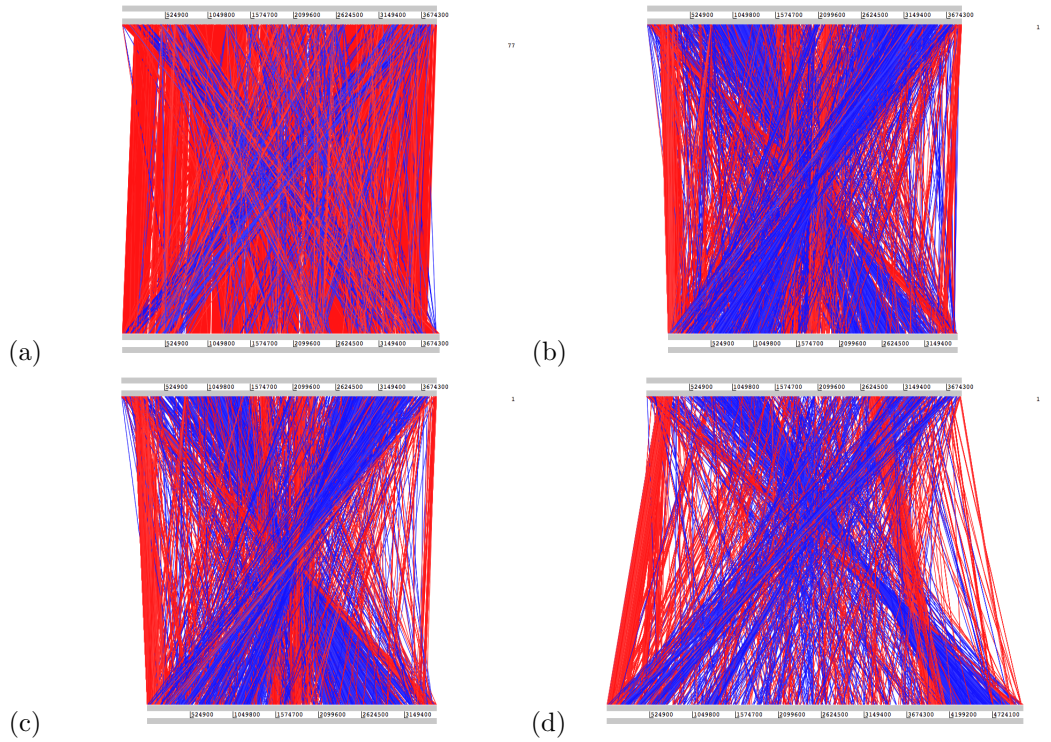


Figure 2.7: Genome rearrangement comparison between rearranged chromosome sequence of *G. thermoglucosidasius* NCIMB 11955 (top sequence) and (a) *G. thermoglucosidasius* C56-YS93, (b) *G. thermodenitrificans* NG80-2, (c) *G. kaustophilus* HTA426, (d) *Bacillus megaterium* DSM319. This comparison was done using ACT software. Red denotes orthologous genes in the same orientation and blue denotes genes in a reverse orientation.

However, comparison between strains in clade "thermoglucosidasius" shows a higher level of rearrangement than observed between the strains of clade "kaustophilus". The largest regions of genome conservation are located at the beginning, end and the middle (800kb) of the genome sequence. The core conserved genes found in the middle of the sequences are responsible for maintaining gene cluster responsible for vitamin B_{12}biosynthesis (discussed in later section), nitrogen metabolism, core genes of the central carbon metabolism, drug resistance as well as encoding elements required for spore formation. By comparison the fermentative pathway is located in the genome of *Geobacillus thermoglucosidasius* NCIMB 11955 starting from 3284141bp from the origin of replication, in the middle of the second genome rearrangement segment. The conservation observed within these three regions might reflect the gene set core to the genus *Geobacillus* and the variable regions correspond to the unique features of individual strains.

Interestingly, the difference between gene order can be found within the genomes of substrains of *G. thermoglucosidasius*:namely C56-YS93 and 11955. This strain variation is observed at the level of metabolic adaptation to environment (discussed in detail in the next section)

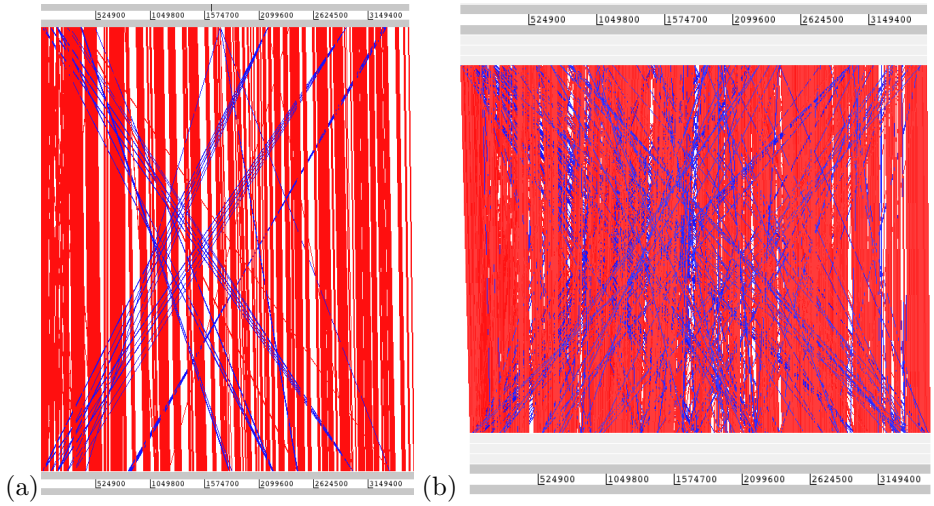


Figure 2.8: Genome arrangement comparison between:(a) *G. kaustophilus* HTA426 vs *G. thermoleovorans* CCB-US.3-UF5 (b) *G. thermoleovorans* CCB-US.3-UF5 with relation to *G. thermodenitrificans* NG80-2. This comparison was done using ACT software. Red denotes orthologous genes in the same orientation and blue denotes genes in a reverse orientation.

as well as when analysing individual gene sequences. This is shown in Figure 2.10 and 2.9 where, sequence identity was calculated using bidirectional best hit and the colour spectrum on the legend corresponds to sequence identity (purple denotes highest sequence identity and red lowest). The greatest difference in sequence can be observed for N-acetylneuraminate synthase (EC 2.5.1.56), Glycerol-3-phosphate transporter, xylose transporter and galactose transporter. Sequences of aldose1-1 epimerase, alcohol dehydrogenase, thioredoxin or beta-xylosidase.

The analysis also included the function based comparison where metabolic subsystems were examined with respect to all the elements needed in a functional subsystem. Compared to other selected strains; *G. thermoglucosidasius* NCIMB 11955 is equipped with the fully functioning: D-sorbitol and L-sorbose utilization, D-tagatose and galactitol utilization, mannose metabolism, glycerol and glycerol-3-phosphate uptake and utilization, organic sulfur assimilation, D-galactarate, D-glucarate and D-glycerate catabolism, oxidative stress response or zinc resistance subsystems.

The sequence based analysis of the genomes of *G. kaustophilus* HTA426 and *G. thermodenitrificans* CCB-US.3-UF5, shows that the similarity on average spans between 50-80%. This illustrates the discrepancies on a genomic level, and functional level, which reflects unique capabilities of *Geobacillus thermoglucosidasius* NCIMB 11955 (see supplementary data for the function based and sequence based comparison between strains). Although the discrepancies might be attributed to evolutionary distance between the strain, the unique metabolic

profiles of the strains can be attributed to acquisition or loss of genes. This is discussed further in Chapter 3.

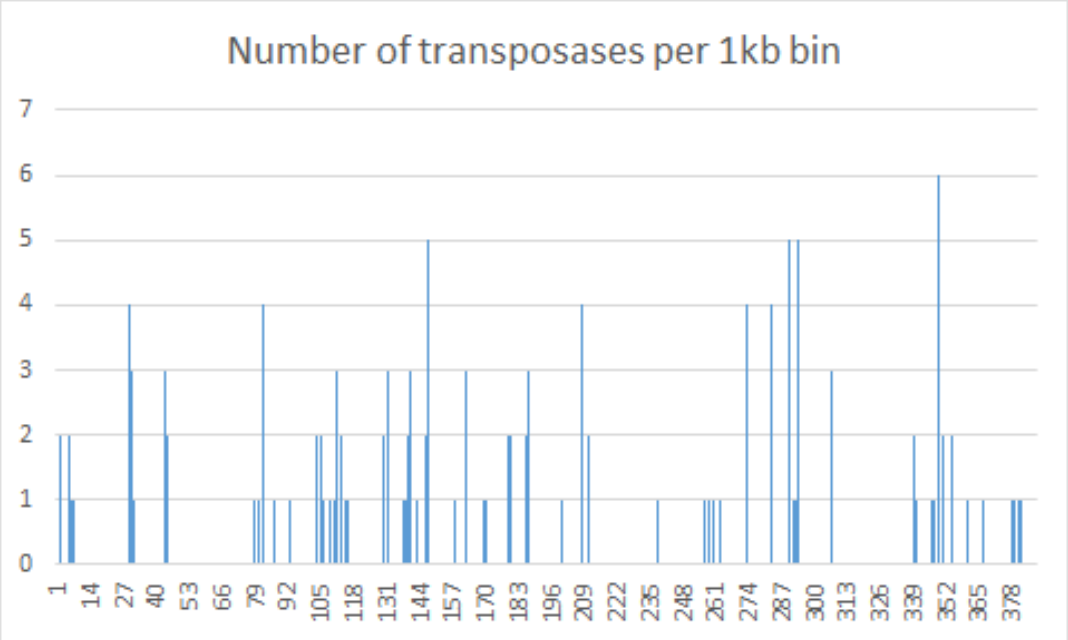


Figure 2.9: The number of transposoases per 1kb in *Geobacillus thermoglucosidasius* NCIMB 11955 .

The comparison of the gene operons in *G. thermodenitrificans* CCB-US.3-UF5 and *G. kaustophilus* HTA426 was performed against the genome of *G. thermoglucosidasius* NCIMB 11955 (see Figure 2.10). When analysing *G. thermoglucosidasius* NCIMB 11955 against *G. kaustophilus* HTA426 , the two strains share majority of the core metabolic elements however strain NCIMB 11955 appears to be differently equipped in several categories making them core differences between the strains (see supplementary data). Clear differences emerge in the categories of amino acids and derivatives, carbohydrates, cofactors, vitamins, prosthetic groups, pigments, fatty acids, lipids and isoprenoids, iron acquisition, membrane transport, nitrogen metabolism, nucleosides and nucleotides, stress response and resistance to antibiotics (see supplementary data). *G. thermodenitrificans* CCB-US.3-UF5 has been found to have unique annotations shared by no other *Geobacillus* strains that play a role in core metabolism. Most notably this strain utilises (NADP) dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9) as an alternative route to catalyse conversion from glyceraldehyde-3-phosphate to glycerate-3-phosphate in one step. The other unique gene annotation for *G. thermodenitrificans* is the presence of pyruvate synthase (EC 1.2.7.1, previously described in the literature [52]) that facilitate conversion of pyruvate to acetyl-CoA in one reaction, which is a tactic unseen in other *Geobacillus* strains, *Bacillus* or even *E.coli*.

It should be noted that above findings are the results of *in silico* analyses and further experimental approach would have been needed to validate our conclusions It can be argued that such differences might be a result of poor annotation. We have however performed comprehensive BLAST searches on the missing genes for both genomes with no significant hits returned (as discussed in the materials and methods section).

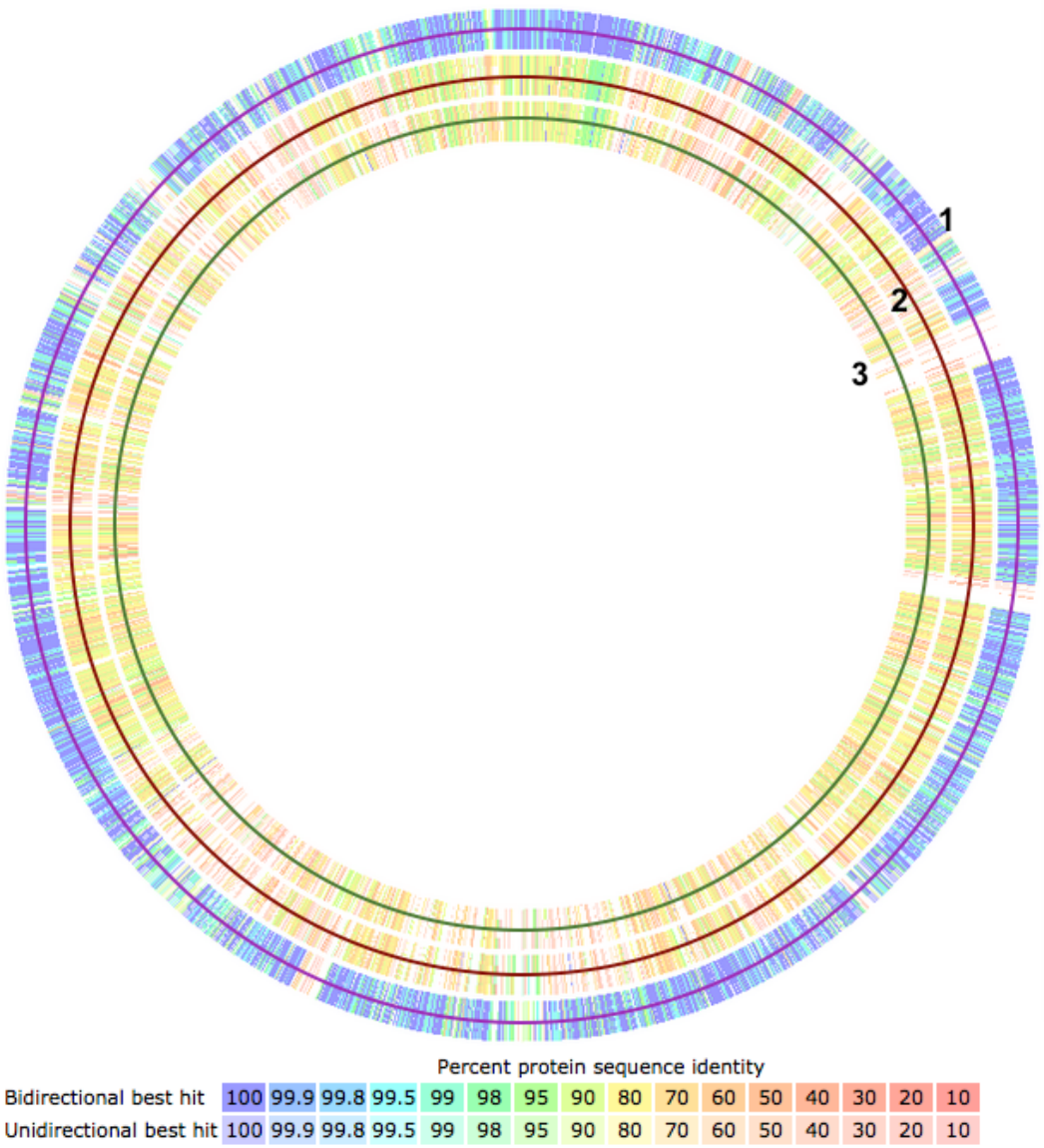


Figure 2.10: Genome sequence comparison between *G. thermoglucosidasius* NCIMB 11955 and: C56-YS93 (1), *G. thermodenitrificans* CCB-US.3-UF5 (2) and *G. kaustophilus* HTA426 (3). The circles are divided into top and bottom part. Top part always relate to *G. thermoglucosidasius* NCIMB 11955 and bottom to strain compared to. The colours denote the results of bi-directional and uni-directional BLAST.

2.3 Hemicellulose Degradation

The ability to degrade the cell walls of plants gives a significant advantage to microorganisms, which can then use this abundant natural feedstock for growth [45, 99]. It is however, a challenging task due to a complex structure of the plant cell wall, which is composed, primarily, of cellulose, hemicellulose, pectin and lignin [45].

While Geobacillus spp do not appear to be cellulose degraders they can produce a battery of enzymes associated with hemicellulose and pectin degradation[45]. The structure of hemicellulose depends on the plant type, but is typically comprised of a backbone of xylan, arabinan (and mixed arabinoxylans), xyloglucan and mannan[45]. The backbones are often branched and decorated by acetylation or methylation [99]. Degradation of hemicelluloses by Geobacillus spp does not require extracellular conversion to monomers, as they use carbohydrate transporters which are able to recognize and transport short oligomers, which may retain their decoration [45]. Therefore, the typical strategy is to produce a limited range of fully secreted enzymes, required for conversion of polymers to oligomers, transport of the oligomers into the cell and further conversion to monomers intracellularly, using an arsenal of endo-1,4-β-xylanases, α-1-arabinofuronosidases, α-glucuronidases acetyl, xylan esterases, β-xylosidases and α-4-O-methyl glucuronidases [45, 99].

2.3.1 Hemicellulose degradation loci in *Geobacillus* spp.

Genes encoding for hemicellulose degradation function have been widely and especially well studied in *Geobacillus stearothermophilus* T-6. It was in this strain that a locus with all the components necessary for hemicellulose degradation (HUS) was first discovered and described.The locus contains thirteen clusters, twelve of which have known and experimentally characterised degradation capabilities. Recently, DeMaayer *et al* 2014 [45] widened the analysis of prevalence of this particular hemicellulose utilization locus to all sequenced *Geobacillus* strains, available at the time (24 strains in total).He reported that, in most strains (clade "thermoglucosidasius" being the exception), these genes can be found between two marker genes; namely the genes encoding enoyl-CoA hydratase (*echD*) and nitropropane dioxygenase (*npd*) (see Figure 2.12 for a view of loci in *Geobacillus* as presented by DeMaayer *et al* 2014). The hemicellulose utilisation locus with its 13 gene clusters has been found to be present in its entirety in only a handful of strains. The majority differed in the number of cluster ranging from 3 to 12, and included incomplete clusters and regions flanked by transposons.

G. thermoglucosidasius CCB_US3_UF5 and B23 have been found to possess only 3 clusters responsible for xylose and arabinose transport and metabolism, suggesting that, these strains simply scavenge for monomeric L-arabinose and D-xylose, rather than degrade hemicellulose polymers/oligomers[45, 105]. Indeed, the study by DeMaayer *et al* 2014, [45] has highlighted a degree of variation in the composition of this locus, which when combined with a large number of transposons (red and black genes in the Figure 2.12) suggests that the locus may be highly mobile.

2.3.2 Hemicellulose degradation in *Geobacillus thermoglucosidasius* NCIMB 11955

Closer inspection of the loci in clade "thermoglucosidasius" (as defined in earlier sections), shows that neither of the strains have hemicellulose degradation genes located between the *echD* and *npd* genes(see Figure 2.13). *G. thermoglucosidasius* C56-YS93, appears to have the hemicellulose degradation genes upstream from the *echD* and *npd* genes, which were defined as the boundary for the locus by DeMaayer *et al*, 2014 and are located between two transposable elements. *Geobacillus thermoglucosidasius* sp. Y4.1MC1, M10EXG and TNO-09.20 do not contain any hemicellulose degrading genes within the defined region (see Figure 2.13) and a study of the region in *G. thermoglucosidasius* NCIMB 11955 shows that these sites are separated from one another by a vast fragment (736 kb) of genome, even though the gene clusters responsible for xylose and arabinose transport and metabolism are present in this strain. The analysis of the position of these gene clusters (xylose and arabinose transport and metabolism), in light with analysis on genome rearrangements between the clades suggests that these genes are not found within the commonly conserved regions and are both found in-between two transposases.

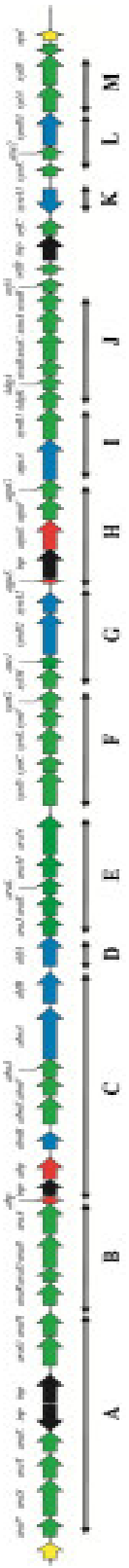


Figure 2.11: Cluser order for hemicellulose degradation locus in *G. stearothermophilus* T-6 Diagram taken from DeMaayer *et al* 2014.

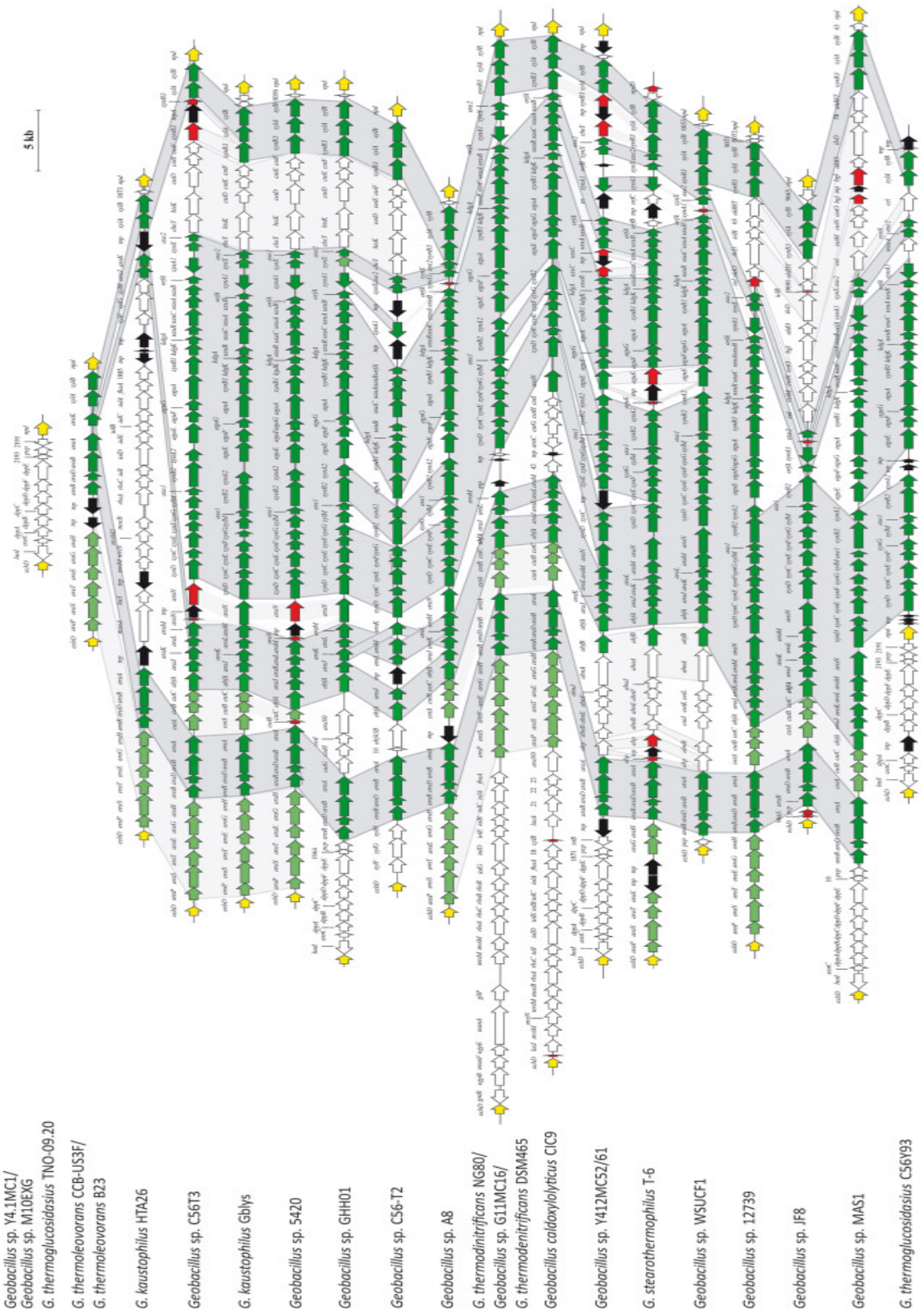


Figure 2.12: Clusters encoding hemicellulose degradation function conservation within genus *Geobacillus*. Black arrows show transposons, and red denotes transposon-disrupted ORFs. The degree of conservation is depicted in the range of green (light green for low conservation, dark green for high). Diagram taken from DeMaayer *et al* 2014.

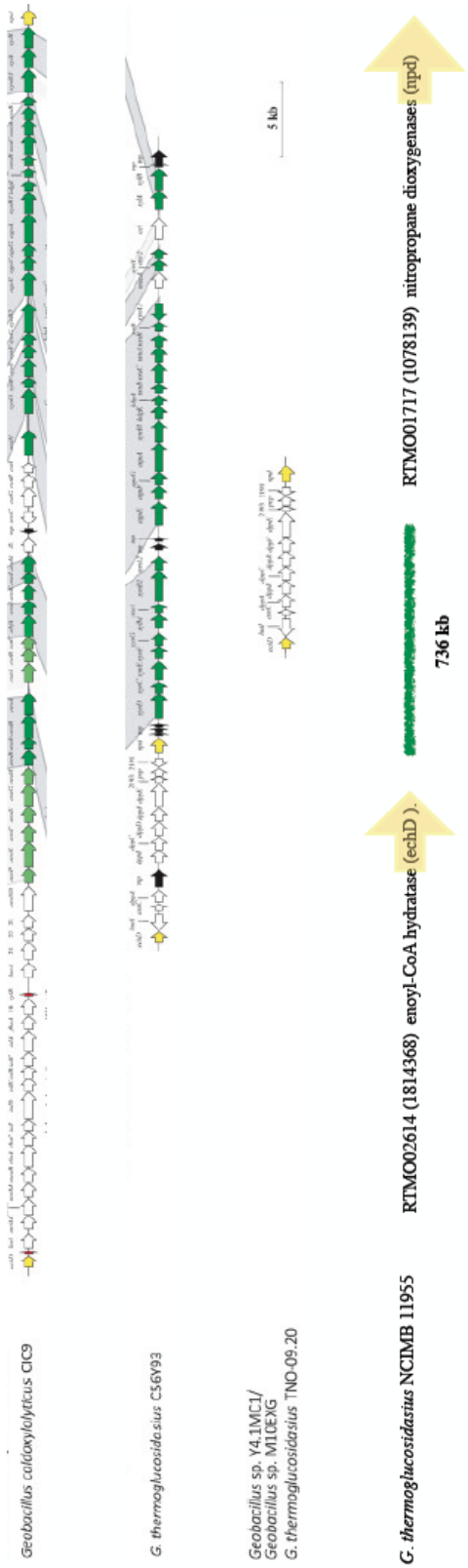


Figure 2.13: Hemicellulose utilization loci in clade "thermoglucosidasius".

2.4 Phage search analysis

The phage search tool (PHAST) [149] was used for identification of prophage sequences within the genomes of: *G. kaustophilus* HTA426, *G. thermoglucosidasius* NCIMB 11955, *G. thermoglucosidasius* C56-YS93, *G. thermoleovorans* CCB-US.3-UF5 and *G. thermodentificans* NG80-2 (Figure 2.14. The consequences of prophage presence can range from regulation and expression of genes and play a significant role in regulation of bacterial population as they interact and coevolve [20]. Phage DNA being a mobile element within the genome can be vector for horizontal gene transfer between bacteria [25]. Prophage sequences are commonly found within the deposited genomes in the NCBI database. However, prophages often remain inactive due to multiple rearrangements, deletions or insertions within its sequence. [25].

Each strain analysed gave a unique profile, with different complete or partial prophage found at different positions on the genomes. Even *Geobacillus thermoglucosidasius* NCBIM 11955 and C56-YS93 differ with respect to position and conserved sequence of the phage with NCIMB 11955 containing only a partial sequence. *G. kaustophilus* HTA426 (Figure 2.14c) is annotated with an intact prophage sequence (integrase, terminase, portal, protease, capsid, head, tail) similar to that of Thermus phage phi OH2 (52.6Kb, 44.60%). "*G. stearothermophilus*" NUB 3621 has two site of incomplete prophage similar to those found in Bacillus phage WBeta (32.1 Kb, GC 41.14%) and Bacillus phage IEBH (34.2 Kb, GC 41.86%)(Figure 2.14). *G. thermodenitrificans* has two sites with incomplete prophage sequences matching: Geobacillus phage GBK2 (19.4 Kb, GC 47.84%) and Thermus phage phi OH2 (24.4 Kb, GC 43.48%). *G. thermoleovorans* CCB US3 UF5 incorporates one intact prophage Thermus phage phi OH2 (45.7 Kb, GC 45.89%) and two incomplete prophage sequences: Bacillus phage G (9.2 Kb, GC 52.12%) and Pithovirus sibericum isolate P1084-T(8.5 Kb, GC 54.68%). *G. thermoglucosiadsius* C56-YS93 has a 55.4Hb region with intact prophage Geobacillus virus E2 (GC 42.91%). Interestingly, *G. thermoglucosiadsius* NCIMB 11955 does not contain any intact prophages, there are however remnants of Bacillus prophage phBC6A52 (23.5 Kb, GC 40.71%). The lack of prophage sequences within the *Geobacillus thermoglucosidasius* can be linked to the high amount of CRISPR elements within its genome. These elements play a role of "bacterial immune system" [110].

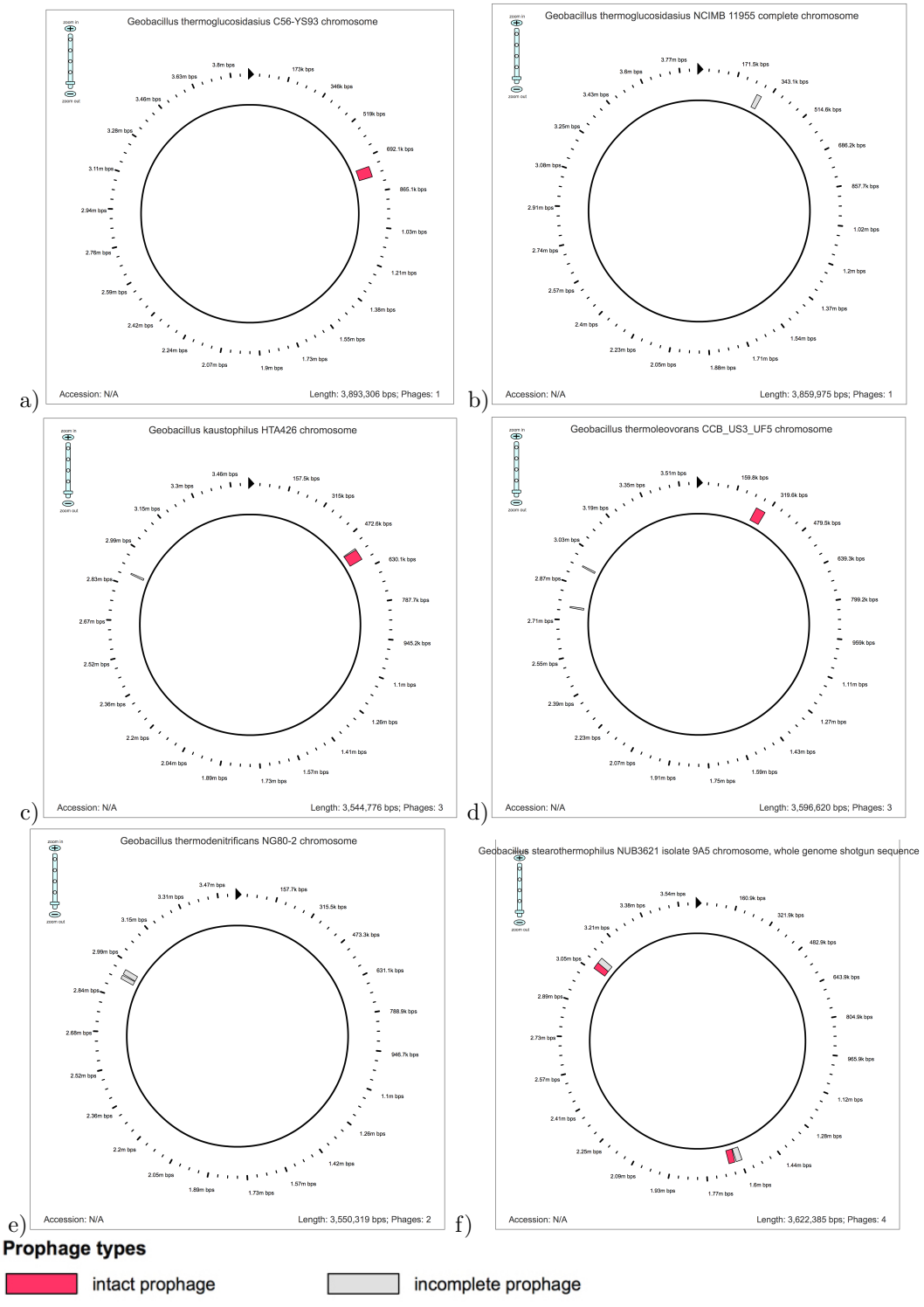


Figure 2.14: Sites of prophage sequence found within the genomes of *G. kaustophilus* HTA426, *G. thermoglucosidasius* C56-YS93, *G. thermoleovorans* CCB US3 UF5, *G. stearothermophilus* NUB3621 and *G. thermodentificans* NG80-2

2.5 Conclusions

This chapter gave an overview of clades within the genus *Geobacillus* and their phylogenetic assignment based on sequence of *recN* gene. The strains can be assigned into five clades although such a distinction does not truly reflect the unique position of each strain within the genus. It could be argued based on the genome rearrangements between type-strains from each clade within the genus, that the genome rearrangement has been observed in all the strains within clade "thermoglucosidasius". This can be supported by the significant number of CRISPR and transposable elements in this clade.

Figure 2.16 shows a phylogenetic tree for *Geobacillus* strains with available proteomes from NCBI Refseq [103, 132, 14] (figure produced and analysis done by Alexander Esin, Imperial College London). It should be noted that as the level of focus increases (from genus to taxa), so does the availability of the sets of orthologous genes, which is further exacerbated by sometimes poor quality of proteomes available in the database. Nonetheless, according to Alexander Esin's diagram (Figure 2.15), *Geobacillus thermoglucosidasius* and it's clade "thermoglucosidasius" have split from the group earlier than the rest of the *Geobacillus* spp. in other groups. This indeed shows this group as truly distinctive from the rest of the strains but complicates the hypothesis proposed in this chapter about the split of the group into strict aerobes and facultative anaerobic microorganisms. Indeed, if the members of the genus have undergone early division into aerobes and facultative anaerobes, strains such as *G. "stearothermophilus"* NUB3621 or *G. caldoxylosyliticus* would contain fermentation genes, conserved within the facultative anaerobes. No such evidence has been found.

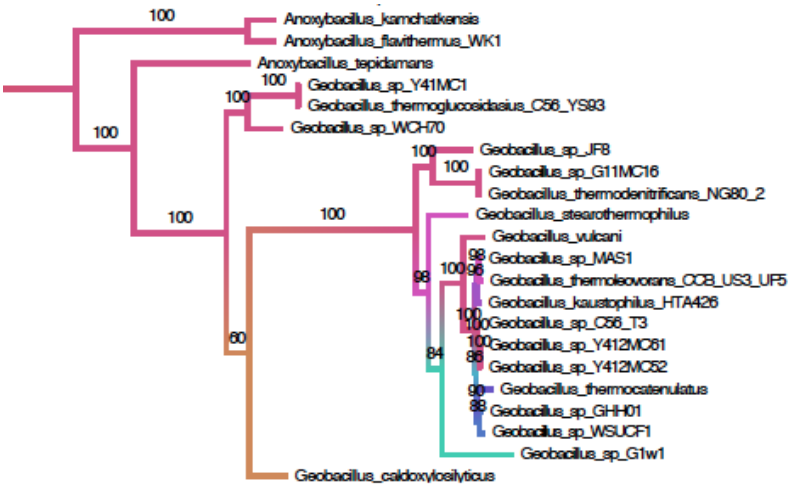


Figure 2.15: Focus on group *Geobacillus* from phylogenetic tree established on super matrix concatenated alignment of orthologous proteins shared by all species in taxa, courtesy of Alexander Esin, Imperial College London.

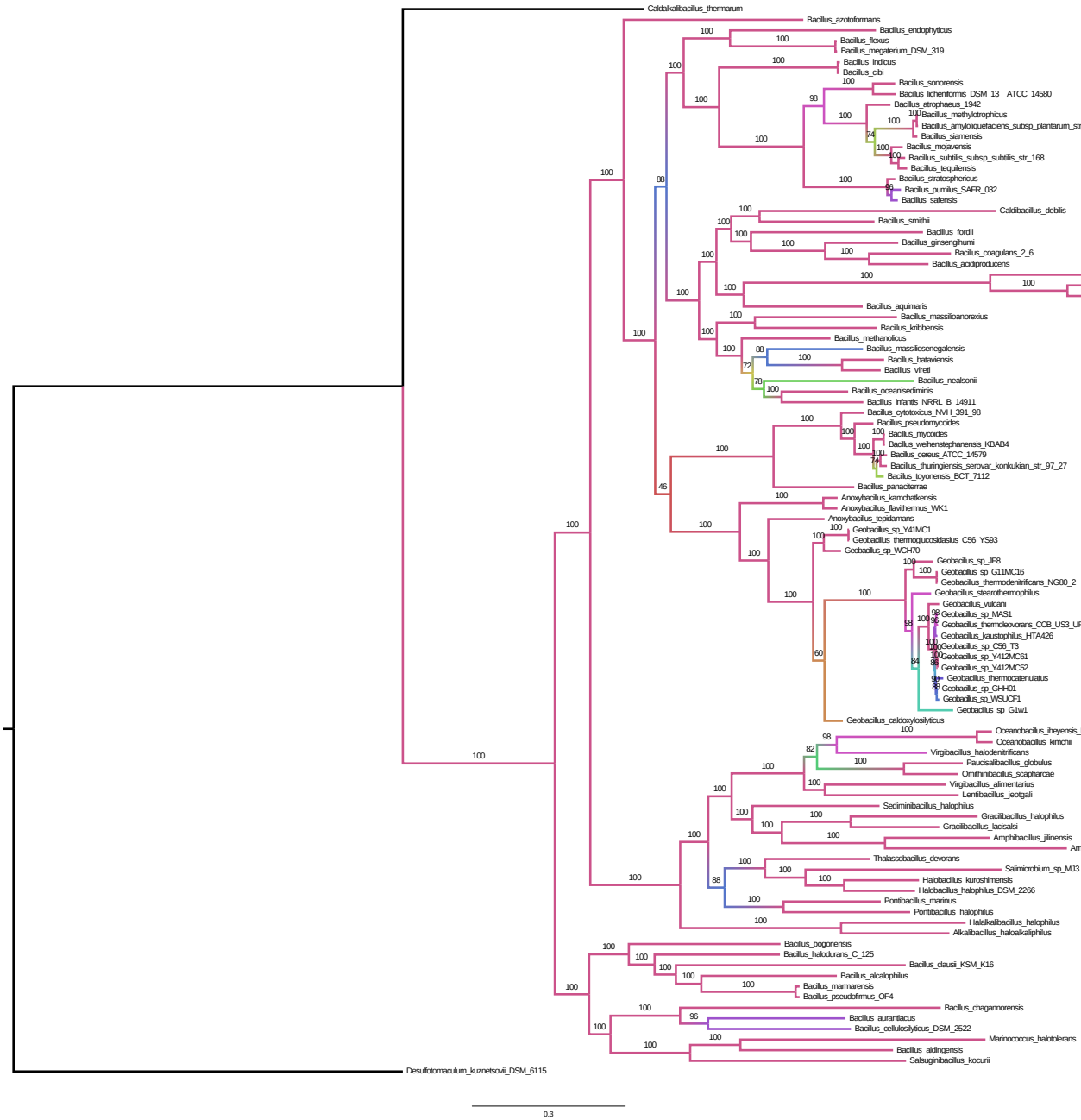


Figure 2.16: Phylogenetic tree established on super matrix concatenated alignment of orthologous proteins shared by all species in taxa, courtesy of Alexander Esin, Imperial College London.The colours of the branches coincide with the bootstrap values. The bootstrap values of 100 are pink.

Chapter 3

PathwayBooster

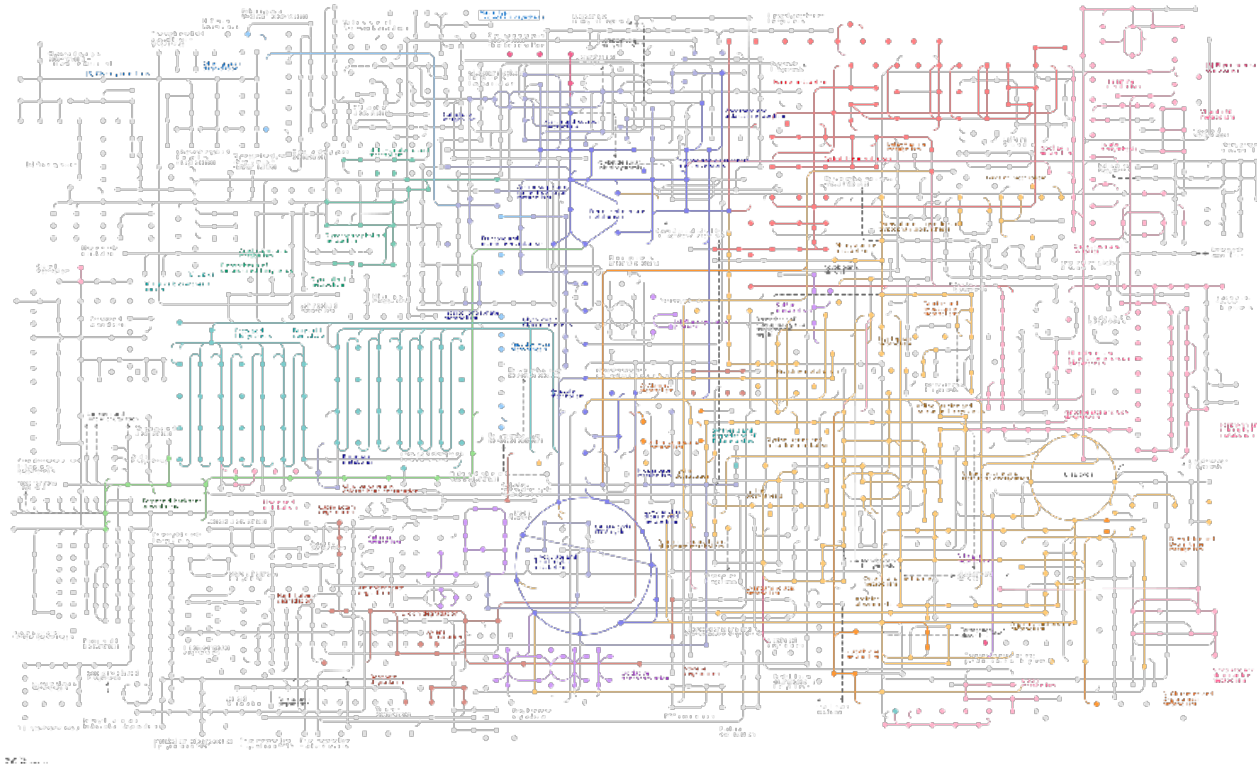


Figure 3.1: A visual representation of genome annotation for *Geobacillus thermoglucosidasius* NCIMB 11955.

3.1 Introduction

The aim of this project was to investigate and discover metabolic capabilities of *Geobacillus thermoglucosidasius* NCIMB 11955. The initial step in this pursuit was to analyse the quality of the gene assignments for this bacterium. At the initial step, the annotation provided by the ERGO Integrated Genomics [104] was investigated. It soon became evident that over half of gene assignments were too general or predicted to encode "hypothetical proteins". In order to widen the possible annotations RAST SEED annotation server [11] was used in order to curate the unspecified gene assignments. The RAST SEED is a commonly used server for annotation of bacterial and archaeal genomes [11]. The annotation is based on gene clusters and protein families that share function and structure, called FIGfams. The assumption driving the gene assignment is that if proteins in a given family are orthologous to one another, they share similar function [11]. The ORFs of unspecified or too general gene annotations were individually manually curated with respect to gene order as well as the assignments generated by both RAST SEED and ERGO Genomics. Based on this methodology, confidence scores were assigned to each ORF (the detailed methodology can be found in Materials and Methods chapter).

It became quickly evident that the comparison between gene assignments generated by annotation servers, might still not be enough to investigate metabolic pathways available for *Geobacillus thermoglucosidasius* NCIMB 11955. The problematic assignments were those, which assigned genes to a protein family rather than to a specific enzyme and the enzymes with unique and unusual annotations for the strain. It became clear that the gene assignments with regards to metabolic pathways should be compared to that of other strains within genus *Geobacillus*. It was hence necessary to create a tool, which could assist this route of investigation. PathwayBooster was developed for this purpose jointly through a collaboration between University of Bath and Imperial College London by Dr Rodrigo Liberal, Beata Lisowska, Prof. David Leak and Dr John Pinney [74]. The coding and methodology for this tool was developed equally by the author and Dr Rodrigo Liberal. Dr Rodrigo Liberal developed the the graphical user interface (GUI) for PathwayBooster and the corrections after initial submission of the manuscript was done by Dr Rodrigo Liberal. The author was also responsible for the case-study and software testing. All the authors of the paper contributed to drafting of the paper. PathwayBooster deals with the problem of mis-annotations and "holes" (mostly created by too general gene assignments) in the metabolic pathways as illustrated by Figure 3.2 but also gives another layer of information for the assignment of

confidence scores to a gene annotation.

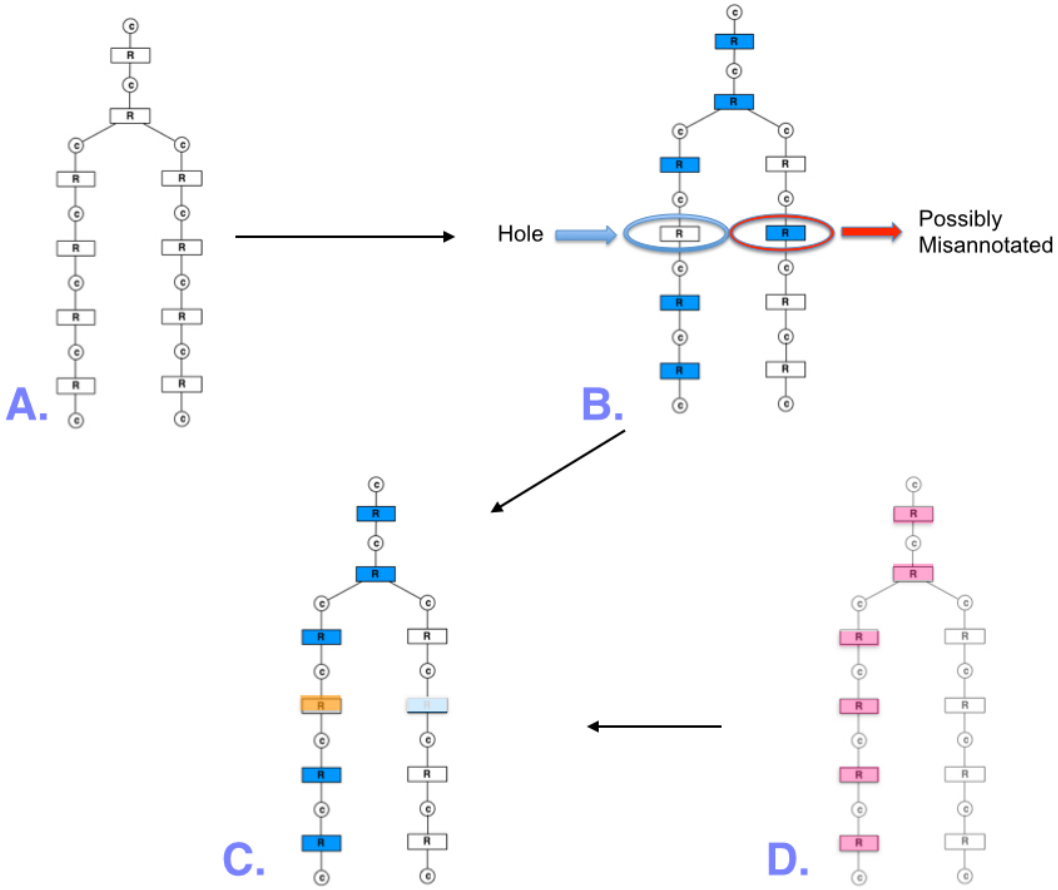


Figure 3.2: The concept behind PathwayBooster. A. shows unannotated pathway, B. annotation stemming from genome annotation, C. reconciled pathway annotation and D. annotation of a pathway in a closely related species. At this stage the reinitiated genes get assigned "confidence scores".

Although other tools are available for addressing the problem of metabolic pathways mis-annotations [93]. PathwayBooster is a unique open-source software tool that was created for the purpose of this project(with Imperial College London [74]) in that it can be especially useful for comparison of annotated metabolic capabilities for multiple closely related species and hence it highlights unique features of a given strain. As illustrated by Figure 3.2, PathwayBooster allows reconciliation of gene annotation for a given metabolic pathway using bi-directional best BLAST (Figure 3.3 b,c), where the candidate for a missing gene is found using a relevant gene from a closely related bacterial strain.

The presence or absence of a given enzymatic assignment for a pathway can be confirmed or refuted using a literature reference search from BRENDA [114, 113, 112](Figure 3.3 d). The results are presented in a graphical form based on KEGG pathways diagrams where up to 7 gene assignments can be visually presented. The annotations of pathways are hyperlinked to information, such as 'Gene Annotations','BLAST-Bidirectional Hits', 'BLAST-3 best hits', 'Literature-Positive', 'Literature-negative', 'Heat Map' and 'XML file' (see Figure 3.3 a-e). A Hamming distance [56, 18] heat map allows a visual assessment of gene conservation between the genomes selected for the analyses of a given pathway (see Figure 3.3 e). These were invaluable in the investigation of annotation of metabolic pathways. In this chapter an overview of metabolism for the members of the genus *Geobacillus* with special regard to *Geobacillus thermoglucosidasius* NCIMB 11955 is presented. The diagrams presented are generated by PathwayBooster. The chapter starts with annotations, which bring this bacterium unique metabolic profile and moves on into discussing other observations that resulted from the whole-genome annotation analysis for this bacterial strain.

a)

1.1.1.1	G_thermoglucosidasius	Geoth_3108	Gt_KEGG
		Geoth_3823	Gt_KEGG
		Geoth_3897	Gt_KEGG
		Geoth_1917	Gt_KEGG
	G_kaustophilus	GK2774	Gk_KEGG
		GK0938	Gk_KEGG
		GK0731	Gk_KEGG
	G_thermodenitrificans	GTNG_1754	Gtn_KEGG
	G_WCH70	GWCH70_0885	Gw_KEGG
	G_Y412MC61	GYMC61_0742	Gy_KEGG

b)

EC Number	Target Species	Target gene	Query gene	Query gene function	EC Number	Seq. similarity	e-value	blast score
	Gt_Rast	fig|6666666.10209.peg.2230	RTMO02626	## Alcohol dehydrogenase II-Gt_GB_Ergo ## Alcohol dehydrogenase II (EC 1.1.1.1)-Gt_Embl_Ergo	1.1.1.1	100.00	0.0	776
		fig|6666666.10209.peg.1784	RTMO03503	## Hypothetical protein-Gt_GB_Ergo ## Hypothetical protein-Gt_Embl_Ergo		100.00	7e-43	164
		fig|6666666.10209.peg.3884	RTMO01820	## Alcohol dehydrogenase-Gt_GB_Ergo ## Alcohol dehydrogenase (EC 1.1.1.1)-Gt_Embl_Ergo	1.1.1.1	100.00	0.0	684
		fig|6666666.10209.peg.1092	RTMO01391	## Alcohol dehydrogenase-Gt_GB_Ergo ## Alcohol dehydrogenase (EC 1.1.1.1)-Gt_Embl_Ergo	1.1.1.1	100.00	0.0	667
		fig|6666666.10209.peg.3881	RTMO01824	## Alcohol dehydrogenase II-Gt_GB_Ergo ## Alcohol dehydrogenase II (EC 1.1.1.1)-Gt_Embl_Ergo	1.1.1.1	100.00	0.0	783
		fig|6666666.10209.peg.4183	RTMO00469	## Alcohol dehydrogenase II-Gt_GB_Ergo ## Alcohol dehydrogenase II (EC 1.1.1.1)-Gt_Embl_Ergo	1.1.1.1	100.00	0.0	783
		fig|6666666.10209.peg.1429	RTMO01758	## 1,3-propanediol dehydrogenase-Gt_GB_Ergo ## 1,3-propanediol dehydrogenase (EC 1.1.1.202)-Gt_Embl_Ergo	1.1.1.202	100.00	0.0	795
		fig|6666666.10209.peg.1751	RTMO03537	## Alcohol dehydrogenase-Gt_GB_Ergo ## Alcohol dehydrogenase (EC 1.1.1.1)-Gt_Embl_Ergo	1.1.1.1	100.00	0.0	801

c)

EC Number	Target Species	Target gene	Query gene	Query gene function	EC Number	Seq. similarity	e-value	blast score
		fig|6666666.10209.peg.1783	RTMO00469	## Alcohol dehydrogenase II-Gt_GB_Ergo ## Alcohol dehydrogenase II (EC 1.1.1.1)-Gt_Embl_Ergo	1.1.1.1	46.00	1e-09	54.3
			RTMO01758	## 1,3-propanediol dehydrogenase-Gt_GB_Ergo ## 1,3-propanediol dehydrogenase (EC 1.1.1.202)-Gt_Embl_Ergo	1.1.1.202	56.82	5e-09	52.4
			RTMO01646	## 1,3-propanediol dehydrogenase-Gt_GB_Ergo ## 1,3-propanediol dehydrogenase (EC 1.1.1.202)-Gt_Embl_Ergo	1.1.1.202	43.75	2e-08	50.4

d)

EC Number	Species	papers	pubmed number
1.1.1.1	E_coli_pub	Nosova, T.; Jousimies-Somer, H.; Kaihovaara, P.; Jokelainen, K.;Heine, R.; Salaspuro, M.: Characteristics of alcohol dehydrogenases ofcertain aerobic bacteria representing human colonic flora. Alcohol.Clin. Exp. Res. (1997) 21, 489-494.	Pubmed:9161610
	G_thermoglucosidasius_pub	Jeon, Y.J.; Fong, J.C.; Riyanti, E.I.; Neilan, B.A.; Rogers,P.L.; Svenson, C.J.: Heterologous expression of the alcoholdehydrogenase (adhI) gene from Geobacillus thermoglucosidasius strainM10EXG. J. Biotechnol. (2008) 135, 127-133.	Pubmed:18436321
	G_thermodenitrificans_pub	Liu, X.; Dong, Y.; Zhang, J.; Zhang, A.; Wang, L.; Feng, L.: Twonovel metal-independent long-chain alkyl alcohol dehydrogenases fromGeobacillus thermodenitrificans NG80-2. Microbiology (2009) 155,2078-2085.	Pubmed:19383697

e)

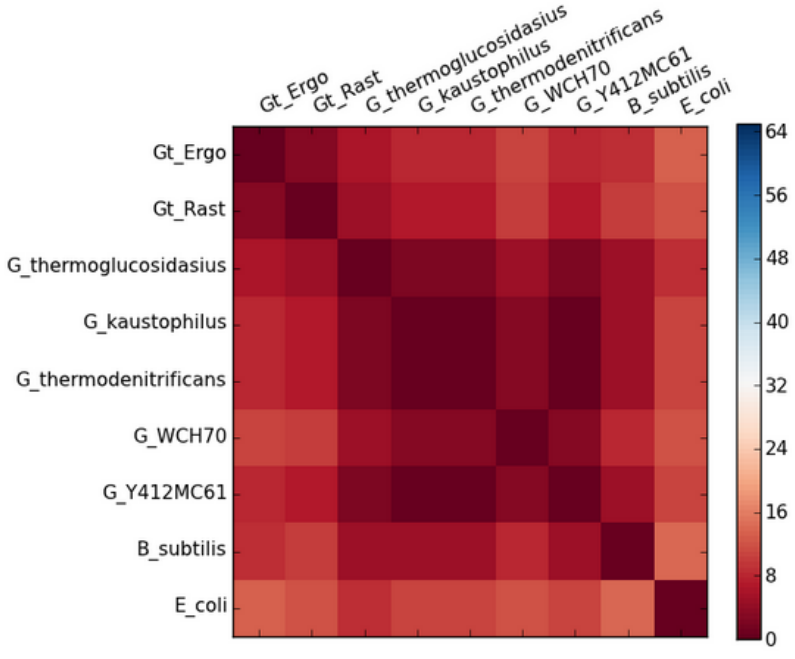


Figure 3.3: Some examples of PathwayBooster output: a) genes corresponding to a given annotation, b)Bi-directional BLAST results, c) three best bi-directional BLASt hits, d) literature search results and e) hamming distance heat map.

3.1.1 Central carbon metabolism

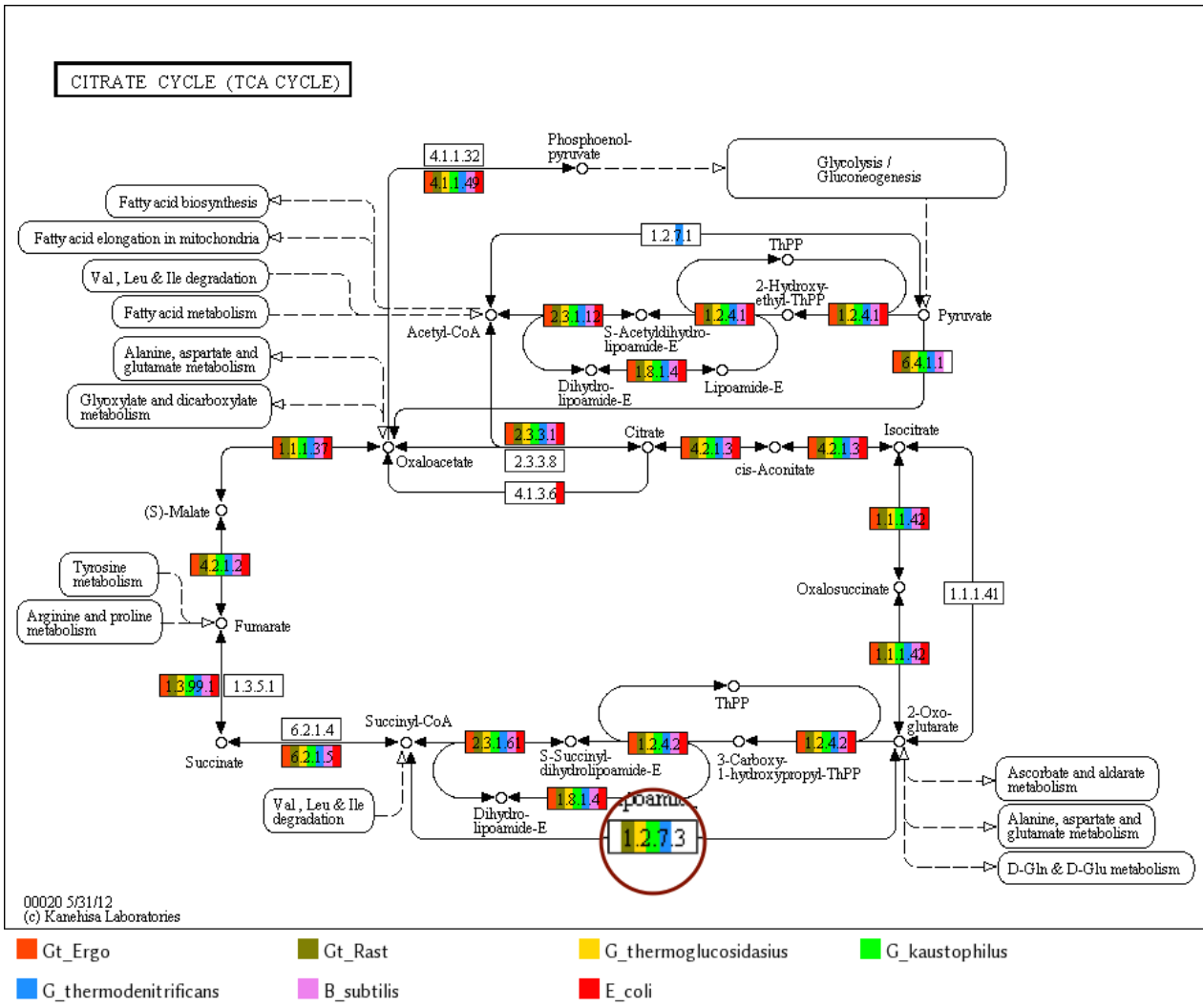


Figure 3.4: PathwayBooster gene annotation for KEGG pathway: TCA cycle. Gt_ERGO and Gt_RAST stand for *G. thermoglucosidasius* NCIMB 11955 annotation done by ERGO Integrated Genomics (Gt_ERGO) and RAST SEED server (Gt_RAST) respectively.

There is a high conservation found within the genus *Geobacillus*. However, *G.thermoglucosidasius* 11955 has been found to have a unique annotation in the glycolytic metabolic pathway with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (NAD(P)+) (phosphorylating) (EC 1.2.1.59). (Figure 3.5). This annotation will be discussed in more detail in the next section, where experimental evidence supporting this annotation is presented. It is worth pointing out a feature shared within the genus of *Geobacillus* which is 2-oxoglutarate synthase (EC 1.2.7.3)(Figure 3.4). This enzyme was first described in the photosynthetic bacterium *Chlorobium thiosulfatophilum* [21] and catalyses reaction (3.1)

$$Succinyl - CoA + CO_2 + Ferredoxin_{red} \rightarrow 2 - oxoglurate + CoA + Ferredoxin_{ox} \quad (3.1)$$

This enzyme is also found and characterised in other thermophilic bacteria and has been shown to be expressed in both aerobic and anaerobic conditions for *Hydrogenobacter thermophilus* TK-6 [128]. No experimental data for this enzyme are found in the literature for any of the members of *Geobacillus* genus.

Members of the genus *Geobacillus*, like their mesophilic counterparts do not possess all the functional elements for a working ED pathway (Figure 3.6) Specifically in *G. thermoglucosidasius* NCIMB 11955 the missing enzyme is 2-dehydro-3-deoxygluconokinase (EC 2.7.1.45) which catalyses the conversion of 2-dehydro-3-deoxy-D-gluconate to 2-dehydro-3-deoxy-6-phospho-D-gluconate (reaction 3.2)

$$ATP + CO_2 + 2-Keto-3-deoxy-D-gluconate \rightarrow ADP + 2-Keto-3-deoxy-6-phosphogluconate \quad (3.2)$$

G. thermoglucosidasius NCIMB 11955 has been found to lack assignment for 2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14). This aldolase precedes step in the ED pathway, which catalyses the reaction (3.3) that leads to the production of D-glyceraldehyde-3-phosphate, which is then used in glycolytic metabolism.

$$2-dehydro-3-deoxy-6-phospho-D-gluconate \rightarrow pyruvate + D-glyceraldehyde3-phosphate \quad (3.3)$$

It is hence assumed that *G. thermoglucosidasius* NCIMB 11955 relies on glycolysis and pentose phosphate pathway for breakdown of glucose to pyruvate.

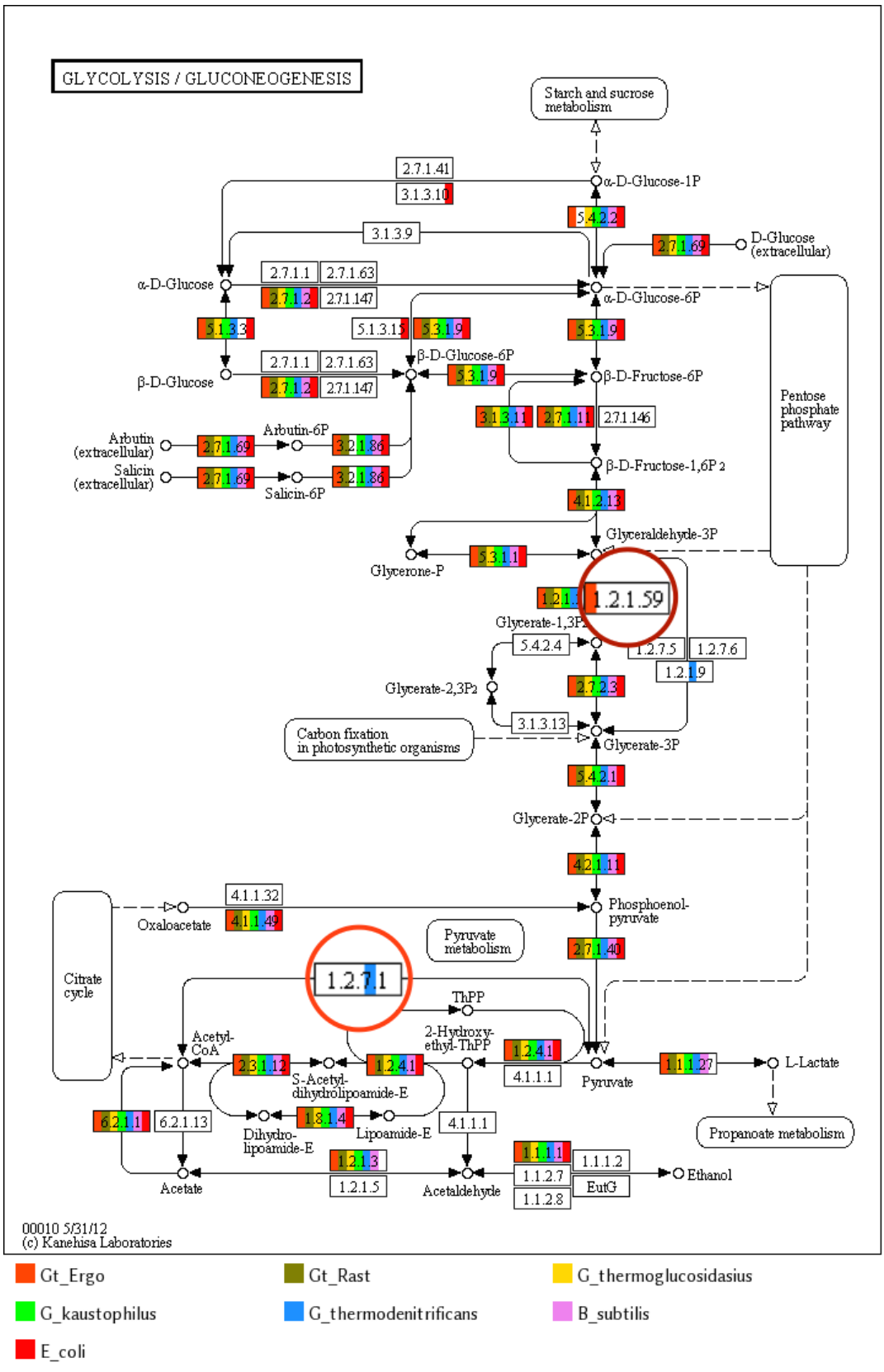


Figure 3.5: PathwayBooster gene annotation for KEGG pathway: Glycolysis and Gluconeogenesis. Gt_ERGO and Gt_RAST stand for *G. thermoglucosidasius* NCIMB 11955 annotation done by ERGO Integrated Genomics (Gt_ERGO) and RAST SEED server (Gt_RAST) respectively.

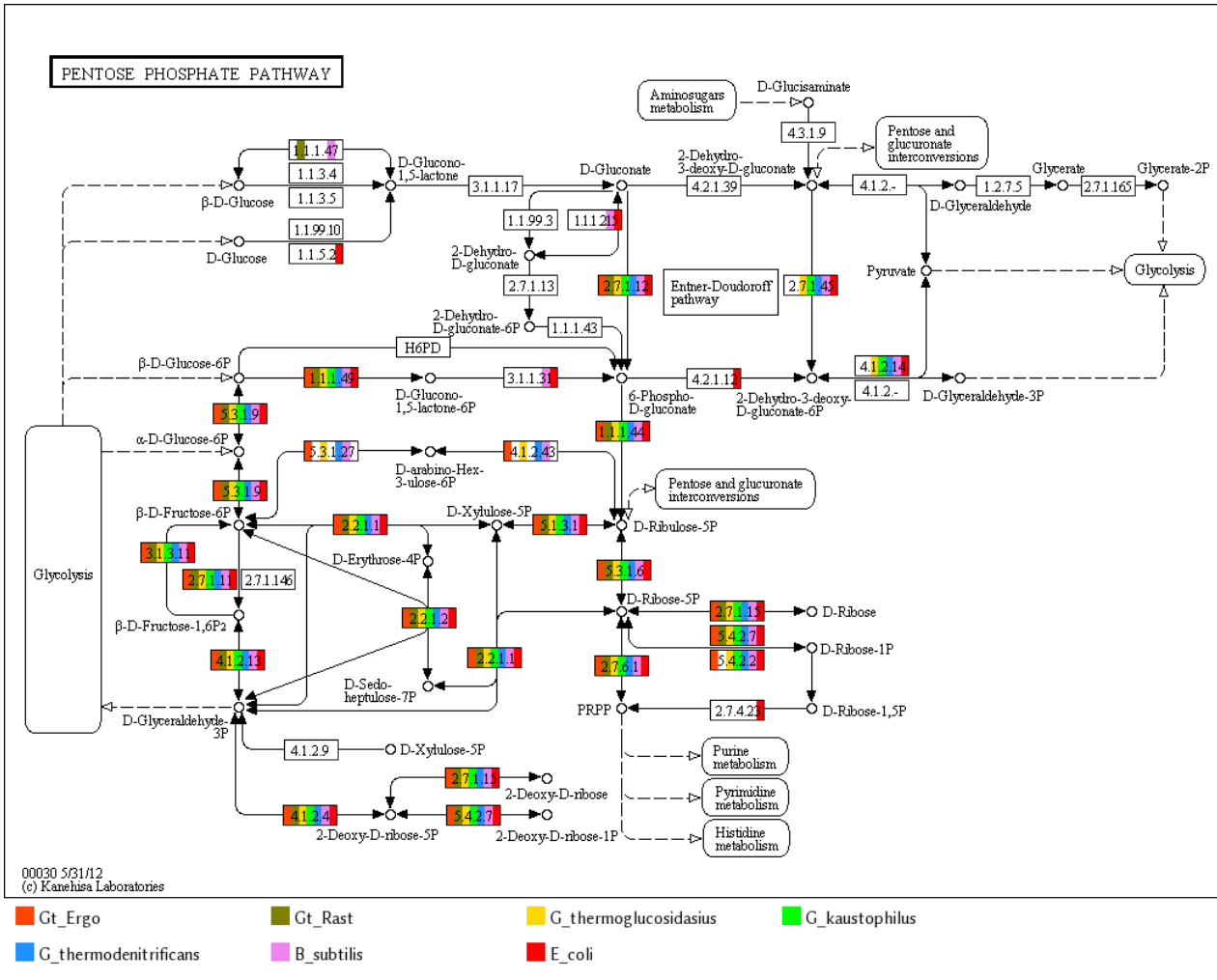


Figure 3.6: PathwayBooster gene assignment for Pentose Phosphate pathway.Gt_ERGO and Gt_RAST stand for *G. thermoglucosidasius* NCIMB 11955 annotation done by ERGO Integrated Genomics (Gt_ERGO) and RAST SEED server (Gt_RAST) respectively.

3.2 Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.59)

3.2.1 Introduction

Description	Max score	Total score	Query cover	E value	Ident	Accession
Geobacillus sp. Y4.1MC1, complete genome	1851	2134	100%	0.0	99%	CP002293.1
Geobacillus thermoglucosidasius C56-YS93, complete genome	1829	2117	100%	0.0	99%	CP002835.1
Geobacillus sp. WCH70, complete genome	1234	1508	99%	0.0	87%	CP001638.1
Geobacillus sp. 12AMOR1, complete genome	884	1144	95%	0.0	80%	CP011832.1
Geobacillus thermoleovorans CCB_US3_UF5, complete genome	870	1149	95%	0.0	80%	CP003125.1
Geobacillus sp. C56-T3, complete genome	870	1149	95%	0.0	80%	CP002050.1
Geobacillus thermodenitrificans NG80-2, complete genome	870	1171	95%	0.0	80%	CP000557.1
Geobacillus sp. LC300, complete genome	866	1154	95%	0.0	80%	CP008903.1
Geobacillus sp. GHH01, complete genome	866	1136	95%	0.0	80%	CP004008.1
Geobacillus kaustophilus HTA426 DNA, complete genome	866	1145	95%	0.0	80%	BA000043.1
Geobacillus sp. Y412MC52, complete genome	861	1144	95%	0.0	79%	CP002442.1
Geobacillus sp. Y412MC61, complete genome	861	1144	95%	0.0	79%	CP001794.1
Geobacillus sp. JF8, complete genome	830	1109	95%	0.0	79%	CP006254.2
Anoxybacillus gonensis strain G2, complete genome	735	970	96%	0.0	76%	CP012152.1
Anoxybacillus flavithermus WK1, complete genome	650	857	96%	0.0	75%	CP000922.1
Bacillus sp. JS, complete genome	578	750	95%	8e-161	74%	CP003492.1
Bacillus toyonensis BCT-7112, complete genome	574	698	93%	1e-159	74%	CP006863.1
Bacillus methanolicus MGA3, complete genome	571	792	96%	1e-158	73%	CP007739.1

Figure 3.7: Nucleotide BLAST analysis results for GAPDH (EC 1.2.1.59).

The analysis of the genome annotation for *Geobacillus thermoglucosidasius* NCIMB 11955, as described in previous sections, has highlighted the presence of an NAD(P)-dependent glyceradehyde-3-phosphate dehydrogenase (GAPDH) (EC 1.2.1.59). This type of GAPDH is absent in the genomes of other strains in the genus *Geobacillus*. This enzyme extends bacterial metabolic capabilities in glycolysis catalysing the reversible reaction as shown in reaction 3.4.

$$GAP + NAD(P)^{+} + P_{\mathrm{i}} \rightarrow 1,3BPG + NAD(P)H + H^{+} \tag{3.4}$$

In the above reaction GAP stands for glyceraldehyde-3-phosphate and 1,3-BPG for 1,3-bisphosphoglycerate.

There are three kinds of glyceraldehyde-3-phosphate dehydrogenases, that can facilitate the conversion from glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate and they can be distinguished by their cofactor-specificity. GAPDH (EC 1.2.1.12), which is the most common within this genus, is an NAD-dependent enzyme and GAPDH (EC 1.2.1.13) utilises $NADP^{+}$.

Description	Max score	Total score	Query cover	E value	N	Accession
Geobacillus sp. Y4.1MC1, complete genome	808	5697	100%	0.0	1	CP002293.1
Geobacillus thermoglucosidasius C56-YS93, complete genome	799	5683	100%	0.0	1	CP002835.1
Geobacillus sp. Y412MC52, complete genome	524	3699	100%	0.0	2	CP002442.1
Geobacillus sp. C56-T3, complete genome	524	3696	100%	0.0	2	CP002050.1
Geobacillus sp. Y412MC61, complete genome	524	3699	100%	0.0	2	CP001794.1
Geobacillus thermoleovorans CCB_US3_UF5, complete genome	520	3706	100%	0.0	2	CP003125.1
Bacillus megaterium QM B1551, complete genome	458	2478	97%	1e-162	2	CP001983.1
Bacillus megaterium DSM319, complete genome	458	2542	97%	2e-162	2	CP001982.1
Bacillus megaterium WSH-002, complete genome	458	2517	97%	2e-162	2	CP003017.1
Bacillus subtilis subsp. spizizenii TU-B-10, complete genome	461	2826	99%	4e-161	2	CP002905.1
Bacillus subtilis subsp. spizizenii str. W23, complete genome	461	2797	99%	5e-161	2	CP002183.1
Bacillus thuringiensis serovar finitimus YBT-020, complete genome	457	2429	98%	3e-160	2	CP002508.1

Figure 3.8: Nucleotide BLAST analysis results for GAPDH (EC 1.2.1.59).

GAPDH (EC 1.2.1.59), which is described in this section, can catalyse the conversion from D-glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate with either NAD^+ or $NADP^+$ as a cofactor [33]. NADP-GAPDHs are found in higher plants, algae and cyanobacteria and facilitates the second step in the photosynthetic carbon reduction cycle [66]. In *Synechococcus* PCC 7942, the $NADP^+$-GAPDH works six-fold more efficiently with $NADP^+$ than with NAD^+ [66].

The annotated GAPDH (EC 1.2.1.59) can be found in cyanobacteria; however the BLAST analysis of the gene sequence suggests that this GAPDH can be found in thirteen other species of the genus (see Figure 3.7). These GAPDH found through the BLAST analysis are assigned to the general family of enzymes (EC 1.2.1.-) and the exact cofactor specificity has not been elucidated. Translated nucleotide BLAST searches have highlighted the presence of this enzyme in only six strains, five of which are present in the clade "thermoglucosidasius" (see Figure 3.8). In this section, the results from expressed, purified and enzymatically assayed GAPDH (EC 1.2.1.59) from *Geobacillus thermoglucosidasius* NCIMB 11955 are presented.

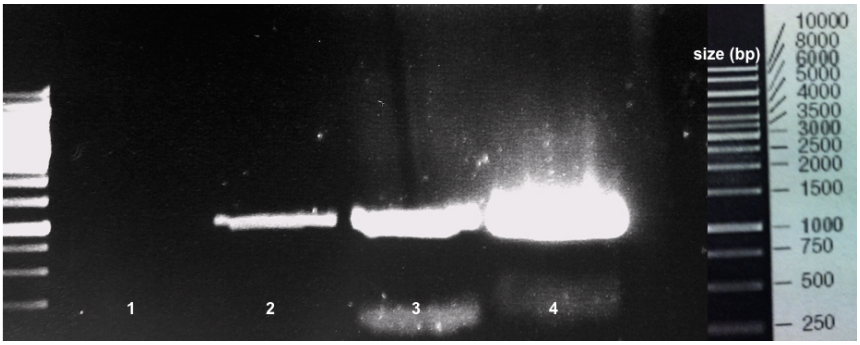


Figure 3.9: Product of GAPDH gene amplification.Lane 2 shows a positive control, lane 1 shows negative control (water), lane 3 and 4 show the amplified GAPDH gene from *G. thermoglucosidasius* NCIMB 11955.

3.2.2 Results

GAPDH cloning and enzyme purification

The GAPDH gene from *G. thermoglucosidasius* NCIMB 11955 was successfully cloned (Figure 3.9) and inserted into pET-28-a vector (Figure 3.10). This work was done firstly by my student Andrew Balfour.

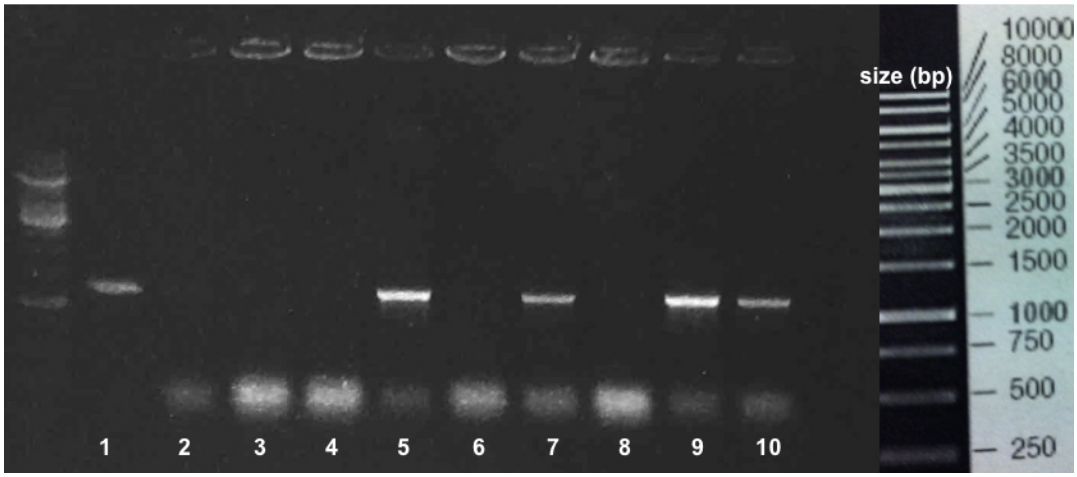


Figure 3.10: Colony PCR showing GAPDH (EC 1.2.1.59) successfully cloned gene. Lane 1 shows a positive control, lane 2 shows negative control, lanes 5,7,9 and 10 show a successfully cloned GAPDH into pET28.

The cell were then induced with a range of IPTG (Figure 3.11) and purified for the enzyme assays.

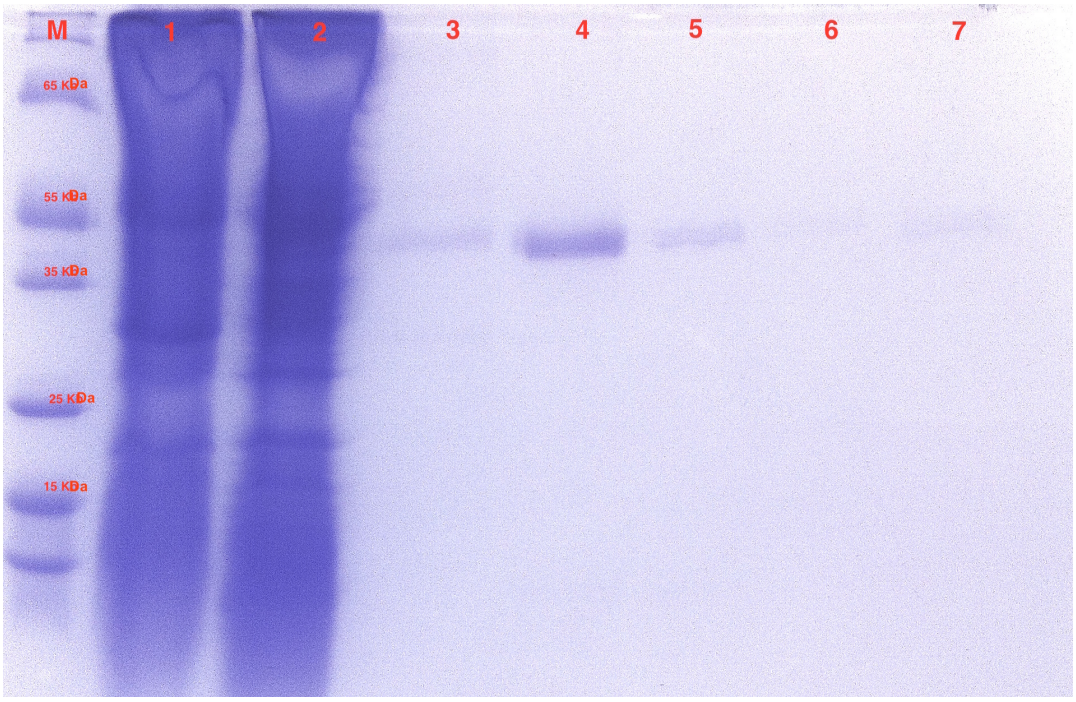


Figure 3.11: SDS-Page gel showing IPTG induced GAPDH (EC 1.2.1.59). M: marker, lane 1: soluble fractions, lane 2: flow through, lane 3: 50 mM , lane 4: 100mM, lane 5: 200mM, lane 6: 500 mM, lane 7: 1M imidazole.The expected size of the GAPDH was 49 kDa.

Enzyme assays.

The reaction efficiency with NAD^+ as sole cofactor is decreased (Michaelis constant is ten fold higher and maximum velocity is three-fold lower), when compared to activity of GAPDH with $NADP^+$ as a sole cofactor. It can be observed that, although GAPDH does utilise NAD^+ as a cofactor, the reaction does not obey the Michaelis-Menten equation. It is speculated from the sigmoidal curve (Figure 3.14, blue) that a level of cooperativity is observed with $NADP^+$. The Hill Equation was used to fit the experimental data (Plot 3.13 and Figure 3.14). The Hill index (h) is 1.6, which suggests positive cooperativity with a cooperativity index (R_a) of 15.58. GAPDH works very efficiently with $NADP^+$ and follows Michaelis-Menten kinetics, which can be observed in Figure 3.15.

The summary of the Michaelis constant and maximum velocity values for both cofactors is shown in Table 3.1.

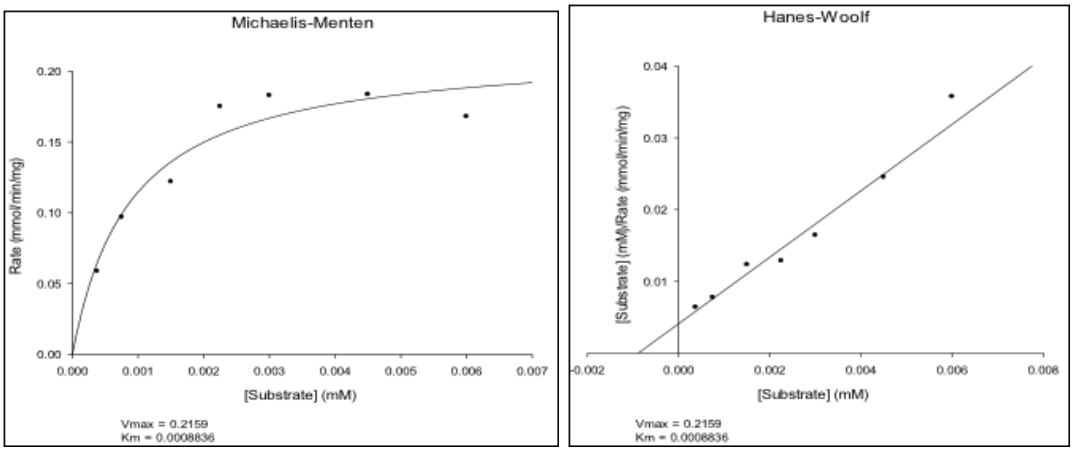


Figure 3.12: Michaelis-Menten and Hanes-Woolf graph for GAPDH with $NADP^+$ as a cofactor and with excess of glyceraldehyde-3-phosphate.

Cofactor	V_{max} (µmol min^{-1} mg^{-1})	K_m / $S_{0.5}$ (µM	k_{cat} (s^{-1})	k_{cat} / K_m (s^{-1} M^{-1})
NAD^+	241	9.1	8.9	0.9 x 10^6
$NADP^+$	727	0.89	26	30.2 x 10^6

Table 3.1: Comparison of derived Michaelis constant and maximum velocity for NAD^+ and $NADP^+$ as cofactors.

3.2.3 Advantages of $NADP^+$/NAD^+-dependent GAPDH.

The presence of an NADP-dependent GAPDH is rare in a non-photosynthetic bacterium [53]. The presence of NADP-dependent GAPDH has been reported in *Bacillus subtilis*[53], however, no such reports have been published for any species in the genus *Geobacillus*. Although efforts have been made to construct a NADP-GAPDH gene knock-in strain in *Geobacillus stearothermophilus* [36], no information on the presence of native NADP-GAPDH has been reported. Our results confirm the presence of NADP-GAPDH in *Geobacillus thermoglucosidasius* NCIMB 11955, and the previous BLAST analysis suggests that the same enzyme can potentially be present in other strains in the clade "thermoglucosidasius", this however has now been reported in the literature and needs extensive laboratory validation. Studies of two GAPDHs (NADP and NAD dependent) in the mesophilic *Bacillus subtilis* as reported by Fillinger *et al.*, 2000 revealed that NADP/NAD-GAPDH is active during gluconeogenesis and repressed during glycolysis [53]. The reported K_m for $NADP^+$ (0.86 µM) and NAD^+ (5.7 µM) GAPDH are similar to those presented in this study, although the K_m for GAPDH from *Geobacillus thermoglucosidasius* is higher (9.1 µM) than that in *Bacillus subtilis* suggesting that the thermophilic GAPDH is more likely to use $NADP^+$ rather than

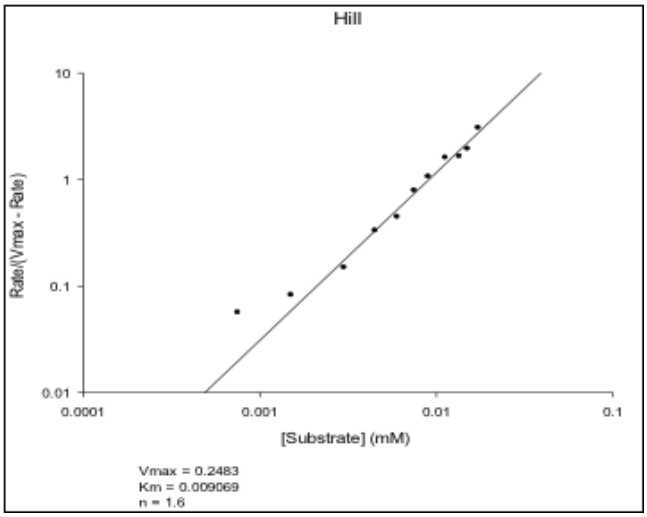


Figure 3.13: Hill graph for GAPDH with NAD^+ as a cofactor. The Hill index (h) is 1.6.

NAD^+.

Since NADPH is an important cofactor in most biosynthetic pathways the presence of NADP-GAPDH provides an important source of this cofactor considering the lack of transhydrogenase across the genus *Geobacillus* (present in *Bacillus subtilis*). As discussed in previous sections, NADPH can be produced via glucose-6-phosphate dehydrogenase (pentose phosphate pathway) (EC 1.1.1.49), which is widely present in *Geobacillus* spp., however, the route is incomplete due to the lack of assignment for 6-phosphogluconolactonase (EC 3.1.1.31). The 6-phosphogluconolactonase catalyses the next step in conversion of glucose-6-phosphate to gluconate-6-phosphate, although, it maybe replaced by a spontaneous reaction [86].

The thermostable NADP-GAPDH from *Geobacillus thermoglucosidasius* NCIBM 11955 can play an important role in the field of cofactor engineering for overall NADP(H) turnover in microbial cell factories, especially those challenged by valuable biochemical production [68]. Although NAD(H) is a major cofactor used in such major pathways such as glycolysis or oxidative phosphorylation and NADP(H) is associated mainly with PPP or amino acid and nucleotide biosynthesis, the effective NADP(H) turnover is essential in genetically-engineered cells challenged with production of complex molecules [68]. The common approach for improving NADP(H) turnover include redirecting fluxes to the NADP-dependent enzymes, and thermostable NADP-GAPDH can be used to that effect alone with over-experession of NADP-GAPDH over NAD-GAPDH.

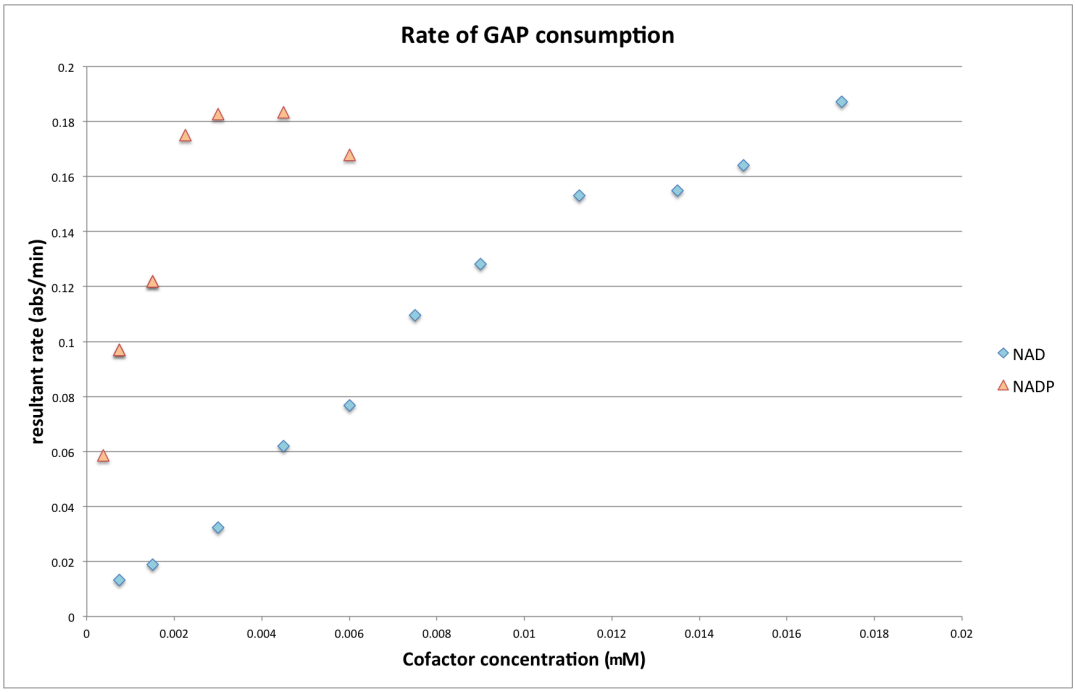


Figure 3.14: Dependence of enzyme velocity on G3P concentration with NAD^+(blue) and $NADP^+$(orange) as a cofactor.

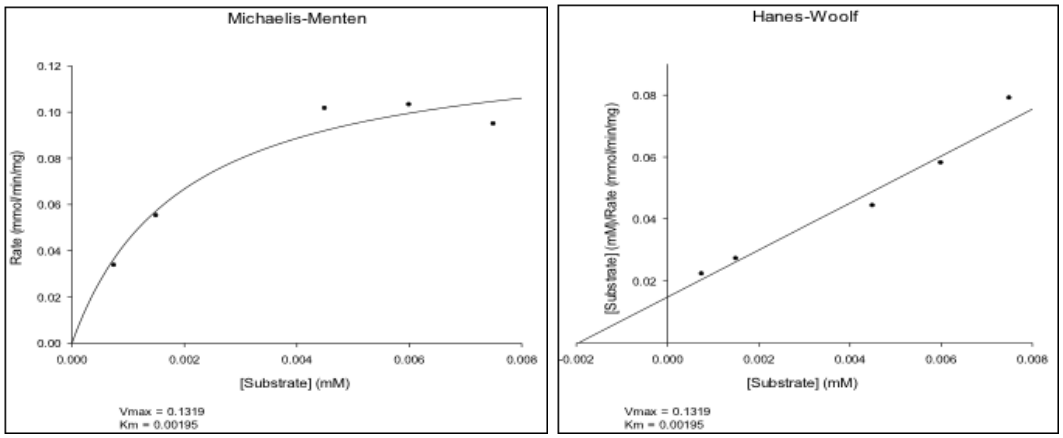


Figure 3.15: Michaelis-Menten and Hanes-Woolf graph for GAPDH as a substrate in excess and $NADP^+$ as a cofactor (marked as substrate on the plots).

3.2.4 Pyruvate metabolism

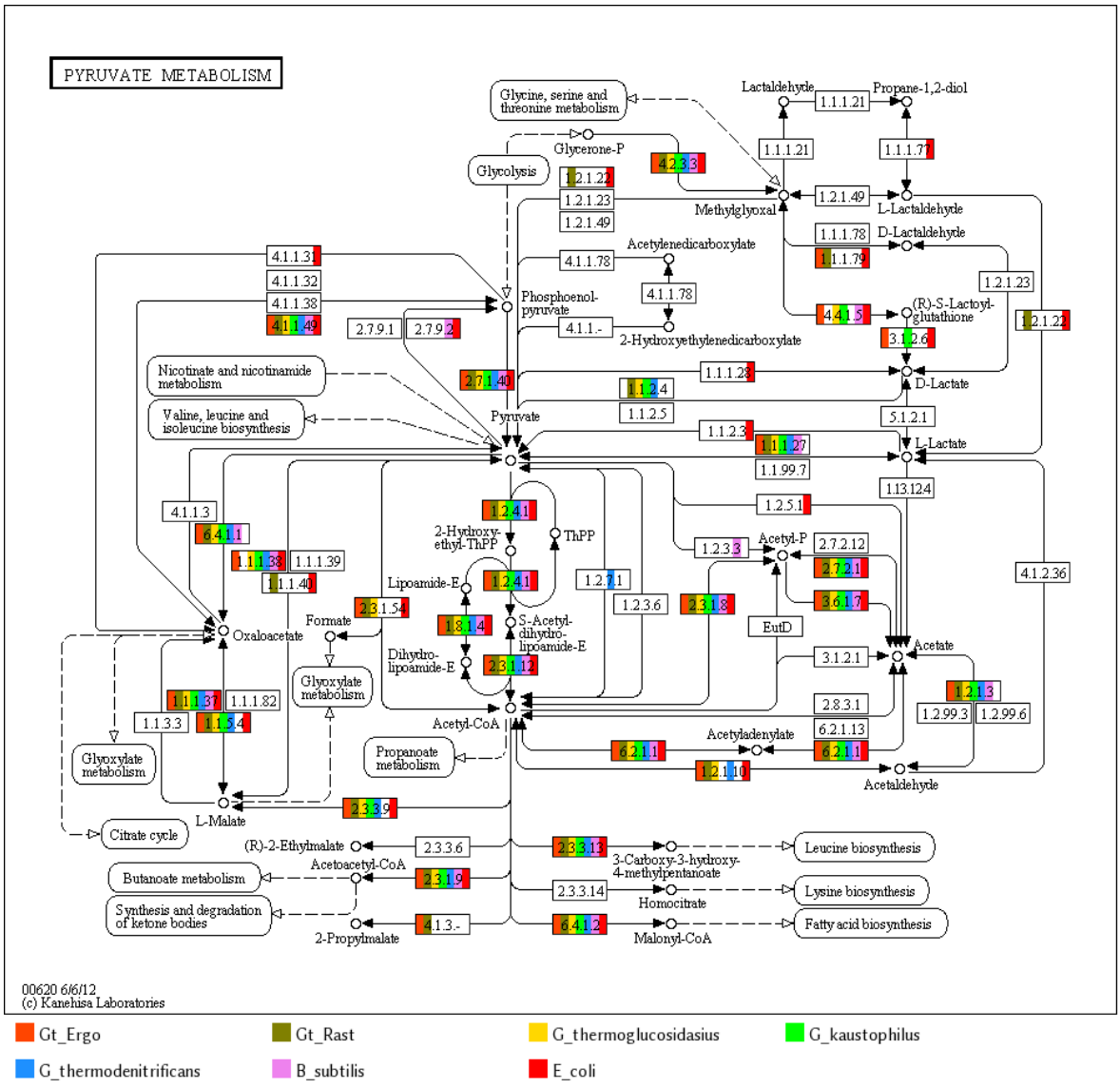


Figure 3.16: PathwayBooster gene annotation for KEGG pathway: Pyruvate Metabolism. Gt_ERGO and Gt_RAST stand for *G. thermoglucosidasius* NCIMB 11955 annotation done by ERGO Integrated Genomics (Gt_ERGO) and RAST SEED server (Gt_RAST) respectively.

Conversion of pyruvate to formate via formate C-acetyltransferase (EC 2.3.1.54) defines a theoretical capability of *G. thermoglucosidasius* NCIMB 11955 and C56-YS93. This transferase can also be found in the genome of Gram-negative *E. coli*.

Furthermore, formate can be formed from oxalate with oxalate decarboxylase (EC 4.1.1.2)

as shown in reaction 3.15.

$$oxalate + H^+ \rightarrow formate + CO_2 \tag{3.5}$$

This is also observed in *Bacillus subtilis* where the manganese containing oxalate decarboxylase requires oxygen for the reaction even though no redox net change is observed ([131, 63]). This explains the presence of a formate efflux transporter that is unique to the strain and highlights the unique catabolic capabilities of *G. thermoglucosidasius* NCIMB 11955 (see earlier section and Figure 3.17).

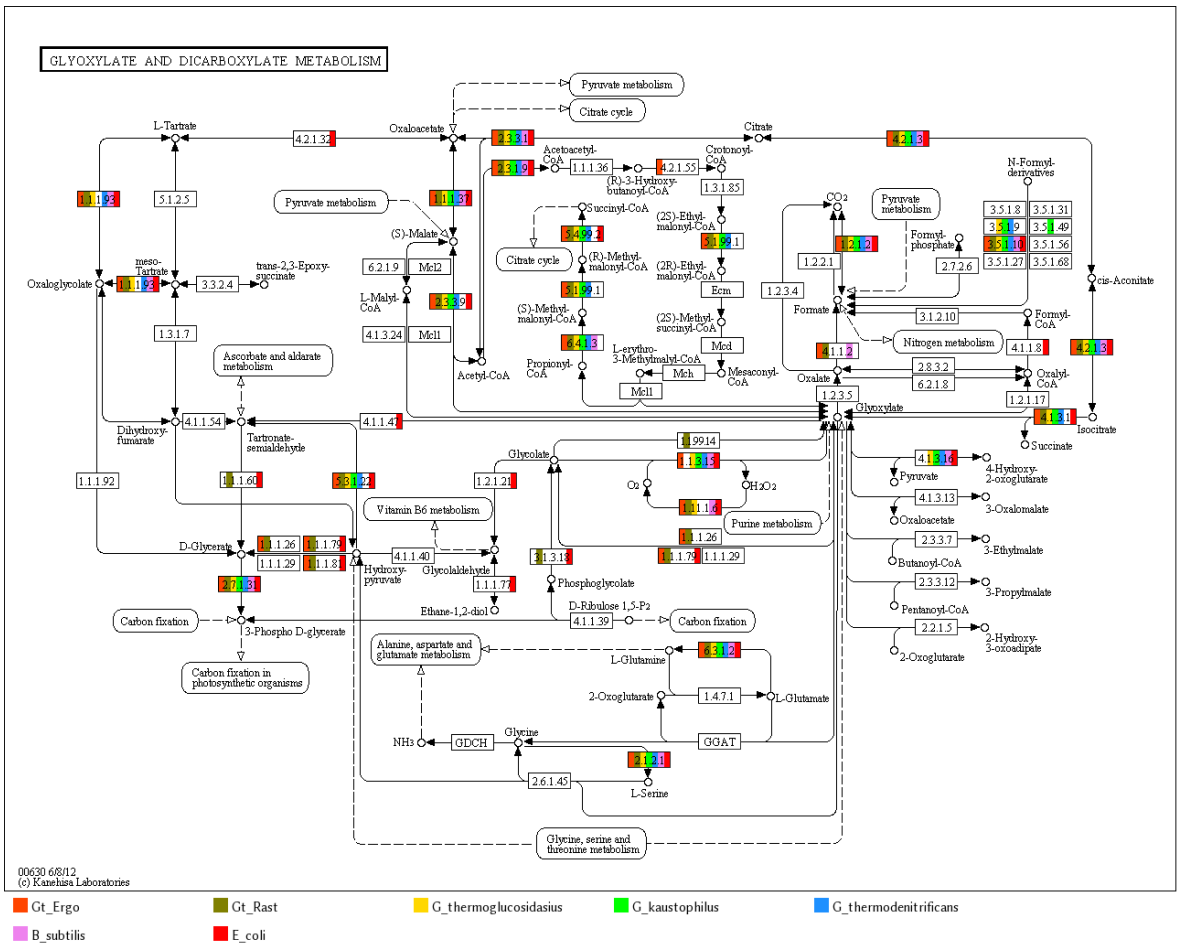


Figure 3.17: PathwayBooster gene annotation for KEGG pathway: Glyoxylate and dicarboxylate metabolism. Gt_ERGO and Gt_RAST stand for *G. thermoglucosidasius* NCIMB 11955 annotation done by ERGO Integrated Genomics (Gt_ERGO) and RAST SEED server (Gt_RAST) respectively.

Methylglyoxal can be catalysed into D-lactaldehyde using glyoxylate reductase ($NADP^+$) (EC 1.1.1.79) in *G. thermoglucosidasius*; however the absence of 2-oxoaldehyde dehydrogenase (NAD^+) (EC 1.2.1.23) is a limiting step in the subsequent conversion of D-lactaldehyde to D-

lactate. Even though glyoxylate reductase is found in *G. thermoglucosidasius* NCIMB 11955 only, 2-oxoaldehyde dehydrogenase is present neither in the genus of *Geobacillus* or *Bacillus* (Figure 3.17). It should be noted that within the gene annotation a second glycoxylate reductase (NAD^+) (EC 1.1.1.26) can be found along with hydroxypyruvate reductase (EC 1.1.1.81). These metabolic capabilities are unique to *G. thermoglucosidasius* NCIMB 11955 and cannot be found in other similar strains.

3.2.5 Pentose and Glucuronate Interconversions

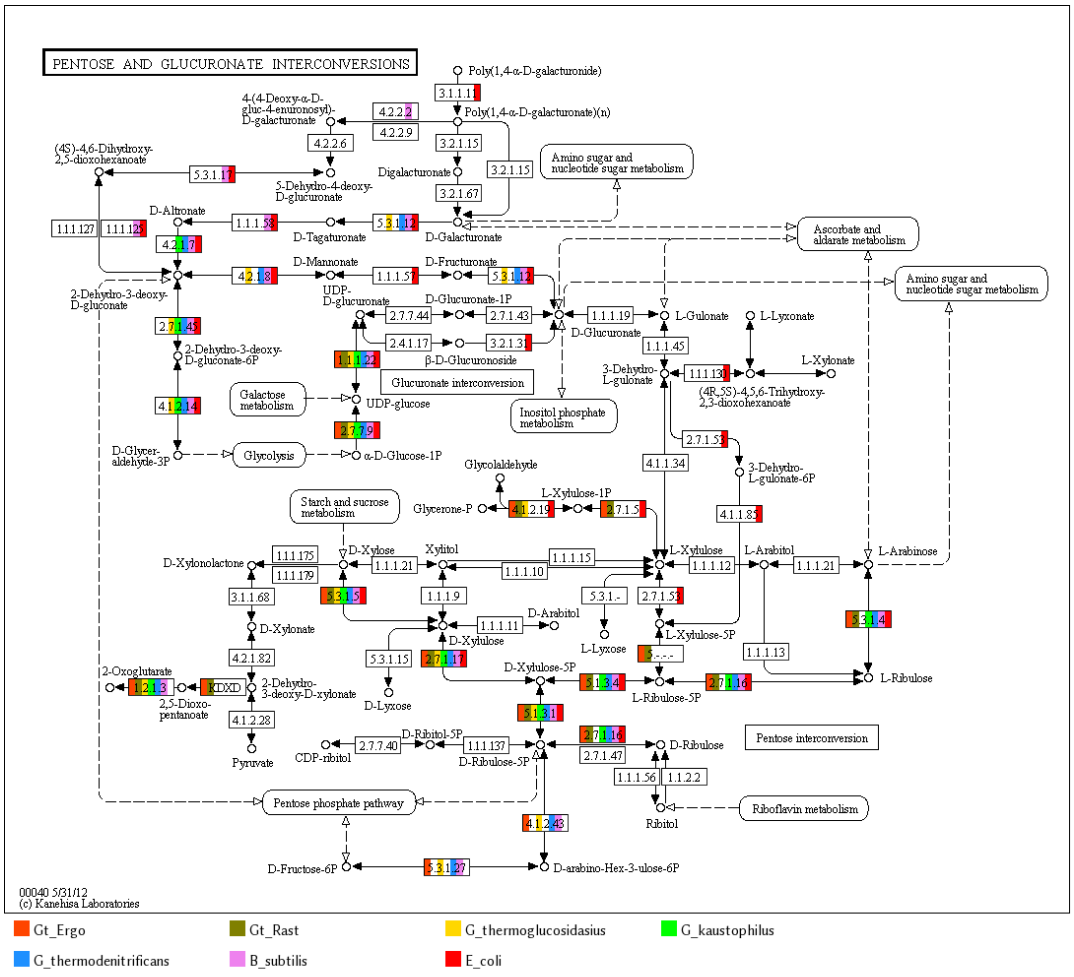


Figure 3.18: PathwayBooster gene annotation for Pentose and Glucuronate interconversions.Gt_ERGO and Gt_RAST stand for *G. thermoglucosidasius* NCIMB 11955 annotation done by ERGO Integrated Genomics (Gt_ERGO) and RAST SEED server (Gt_RAST) respectively.

In the pentose and glucuronate interconversions pathway *G. thermoglucosidasius* NCIMB 11955 has no annotated genes for conversion of D-altronate to D-glyceraldehyde-3-phosphate and pyruvate. These steps involve the following consecutive enzymes: altronate dehydratase (EC 4.2.1.7) [120], 2-dehydro-3-deoxygluconokinase (EC 2.7.1.45) [40] and 2-dehydro-3-deoxy-phosphogluconate aldolase (EC 4.1.2.14) [84]. All three enzymes are present both in all other species in the genus *Geobacillus* as well as in mesophilic *Bacillus* with the exception of *G. thermoglucosidasius* NCIMB 11955 and *Geobacillus* sp. WCH70. Bidirectional BLAST search against the *G. thermoglucosidasius* NCIMB 11955 against altronate dehydratase, 2-dehydro-3-deoxygluconokinase and 2-dehydro-3-deoxy-phosphogluconate al-

dolase rendered no significant results (Figure 3.19 and Figure3.20). Another missing enzyme for the degradation of altronate is altronate hydrolase (EC 4.2.1.8) which catabolyses conversion of 2-dehydro-3-deoxy-D-gluconate to D-mannonate. Altronate dehydratase is also not found in the genome annotation of *Geobacillus* sp. Y412MC61. *G. thermoglucosidasius* NCIMB 11955 appears to also lack gene annotation for glucuronate isomerase (EC 5.3.1.12) [8], which converts D-glucuronate to D-fructuronate and D-glucuronate to D-tagaturonate (Figure 3.18). These enzymes play a major role in the glucuronic acid utilization [8, 84, 40, 120, 117]. In *Geobacillus stearothermophilus* T-6, Shulami *et al.* was first to report that the glucuronic acid utilisation operon plays a significant role in the degradation of xylan. In this bacterium, xylan backbone is firstly cleaved by extracellular xylanases producing xylobiose and oligoxylose units (with side chains such as glucuronic acid) which, after transport inside the cell by specialised permeases α-D-glucuronosyl-xylotriose is degraded inside the cell by the α-glucuronidase to xylotriose and D-glucuronic acid [117] . The D-glucuronate is firstly converted to 2-dehydro-deoxy-D-gluconate (via a three step route) and then to glyceraldehyde-3-phosphate and pyruvate (reaction 3.2 and 3.3). The three-step route catalysing the first breakdown of D-glucuronate starts with the conversion of D-glucuronate to D-fructuronate by glucuronate isomerase (EC 5.3.1.12), which is consequently catalysed by mannonate oxidoreductase to D-mannonate (EC 1.1.1.57) and altronate hydrolase (EC 4.2.1.8) [117]. The lack of gene assignment for altronate hydrolase (EC 4.2.1.8), altronate dehydratase (EC 4.2.1.7), 2-dehydro-3-deoxygluconokinase (EC 2.7.1.45) and 2-dehydro-3-deoxy-phosphogluconate aldolase (EC 4.1.2.14) hence leads to a conclusion that *Geobacillus thermoglucosidasius* NCIMB, along with a few other strains in genus *Geobacillus* does not have a functional pathway allowing glucuronic acid metabolism.The loss of those function as compared to the other strains might be a results of genome rearrangements as described in the previous sections.

A common problem with gene functional assignment which is discussed in detail in "Materials and Methods" chapter, is that some genes might be assigned erroneously by the automated annotation softwares [41]. Commonly, a gene can be assigned with an EC number that points to enzyme family rather than to a specific enzyme annotation. For example, EC 4.2.1.- corresponds to a class "hydro-lyases" whilst a true annotation of EC 4.2.1.8 for altronate hydrolase can be missed. As practically illustrated in the previous sections, those metabolic pathway "holes" can be evident when analysing pathways annotated with PathwayBooster. In those instances, a bidirectional BLAST analysis is performed against a gene sequence belonging to

a closely related species. This approach has allowed the curation of multiple mis-annotations in the metabolic pathways.

Although, *G. thermoglucosidasius* NCIMB 11955 is found to be missing certain enzymes in pentose and glucuronate interconversions pathway, it has been found to have unique added *in silico* metabolic capabilities. These include the ability to catalyse the conversion of L-xylulose to glycerone-phosphate and glucolaldehyde through two consecutive steps. These steps are catalysed by rhamnulokinase (EC 2.7.1.5) and L-rhamnulose 1-phosphate aldolase (EC 4.1.2.19). Rhamnulokinase can carry out two metabolic reactions (reaction 3.6 and 3.7), converting either L-xylulose to L-xylulose-1-phosphate or L-rhamnulose to L-rhamnulose-1-phosphate.

$$ATP + L - Xylulose \rightarrow ADP + L - Xylulose - 1 - phosphate \tag{3.6}$$

$$ATP + L - Rhamnulose \rightarrow ADP + L - Rhamnulose - 1 - phosphate \tag{3.7}$$

Similarly, rhamnulose-1-phosphate aldolase can further catalyse either L-xylulose-1-phosphate to glycerone phosphate and glycolaldehyde (reaction 3.8) or L-rhamnulose-1-phosphate to glycerone phosphate and (S)-lactaldehyde (reaction 3.9).

$$L - Xylulose - 1 - phosphate \rightarrow Glyceronephosphate + Glycolaldehyde \tag{3.8}$$

or:

$$L - Rhamnulose - 1 - phosphate \rightarrow Glyceronephosphate + (S) - Lactaldehyde \tag{3.9}$$

Both the rhamnulokinase and rhamnulose-1-phosphate aldolase play an important role in fructose and mannose metabolism (Figure 3.21), allowing *G. thermoglucosidasius* NCIMB 11955 to subsequently convert L-rhamnulose to glyceraldehyde-3-phosphate. During the breakdown of L-rhamnulose 1-phosphate, dihydroxyacetone phosphate (glycerine phosphate) can be used as source of energy or as a starting material in biosynthesis ([28]).This aldolase has been found to be present predominantly in bacterial strains using L-rhamnose as sole carbon source ([28]).

a)

EC Number	Target Species	Target gene	Query gene	Query gene function	EC Number	Seq. similarity	e-value	blast score
4.2.1.7	Gt_Rast							
	G_kaustophilus	GK1959	RTMO01157	## Sigma-54-dependent transcriptional activator-Gt_GB_Ergo ## Sigma-54-dependent transcriptional activator-Gt_Embl_Ergo		50.00	0.44	26.2
			RTMO03013	## PilU-Gt_GB_Ergo ## Pili retraction protein pilU-Gt_Embl_Ergo		33.96	0.51	25.8
		GK1958	RTMO01419	## Sorbitol dehydrogenase-Gt_GB_Ergo ## Sorbitol dehydrogenase (EC 1.1.1.14)-Gt_Embl_Ergo	1.1.1.14	24.24	0.55	25.8
			RTMO04545	## Bacterial Protein Translation Initiation Factor 2 (IF-2)-Gt_GB_Ergo ## Bacterial Protein Translation Initiation Factor 2 (IF-2)-Gt_Embl_Ergo		46.15	2.1	26.6
			RTMO04464	## Transketolase-Gt_GB_Ergo ## Transketolase (EC 2.2.1.1)-Gt_Embl_Ergo	2.2.1.1	26.09	3.1	26.2
			RTMO04651	## Molybdopterin-guanine dinucleotide biosynthesis protein B-Gt_GB_Ergo ## Molybdopterin-guanine dinucleotide biosynthesis protein B-Gt_Embl_Ergo		54.17	4.4	25.8
			RTMO01994	## Exodeoxyribonuclease VII large subunit-Gt_GB_Ergo ## Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)-Gt_Embl_Ergo	3.1.11.6	43.24	1.4	27.3
		GTNG_1858	RTMO04545	## Bacterial Protein Translation Initiation Factor 2 (IF-2)-Gt_GB_Ergo ## Bacterial Protein Translation Initiation Factor 2 (IF-2)-Gt_Embl_Ergo		46.15	2.1	26.6
			RTMO04651	## Molybdopterin-guanine dinucleotide biosynthesis protein B-Gt_GB_Ergo ## Molybdopterin-guanine dinucleotide biosynthesis protein B-Gt_Embl_Ergo		54.17	4.6	25.8
	G_thermodenitrificans							
		GTNG_1859	RTMO02625	## Aldehyde dehydrogenase-Gt_GB_Ergo ## Aldehyde dehydrogenase (EC 1.2.1.3)-Gt_Embl_Ergo	1.2.1.3	30.91	0.15	27.7
			RTMO01157	## Sigma-54-dependent transcriptional activator-Gt_GB_Ergo ## Sigma-54-dependent transcriptional activator-Gt_Embl_Ergo		53.85	0.17	27.3
			RTMO01904	## Lactam utilization protein LAMB-Gt_GB_Ergo ## Lactam utilization protein LAMB-Gt_Embl_Ergo		34.88	0.24	26.9
	G_WCH70							
	G_Y412MC61							
	B_subtilis	BSU12390	RTMO02913	## DNA polymerase III subunit gamma/tau-Gt_GB_Ergo ## DNA polymerase III subunit gamma/tau (EC 2.7.7.7)-Gt_Embl_Ergo	2.7.7.7	32.39	0.25	30.4
			RTMO04263	## Cystathionine beta-lyase-Gt_GB_Ergo ## Cystathionine beta-lyase (EC 4.4.1.8) / Cystathionine gamma-lyase (EC 4.4.1.1)-Gt_Embl_Ergo	4.4.1.8 4.4.1.1	25.00	0.67	28.9
			RTMO00939	## DNA polymerase I-Gt_GB_Ergo ## DNA polymerase I (EC 2.7.7.7)-Gt_Embl_Ergo	2.7.7.7	24.32	1.1	28.1
			RTMO02913	## DNA polymerase III subunit gamma/tau-Gt_GB_Ergo ## DNA polymerase III subunit gamma/tau (EC 2.7.7.7)-Gt_Embl_Ergo	2.7.7.7	32.39	0.24	30.4
	E_coli	b3091	RTMO01051	## N-acetylglucosamine-6-phosphate deacetylase-Gt_GB_Ergo ## N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25)-Gt_Embl_Ergo	3.5.1.25	28.95	0.86	28.5
			RTMO04306	## Aspartate 1-decarboxylase-Gt_GB_Ergo ## Aspartate 1-decarboxylase (EC 4.1.1.11)-Gt_Embl_Ergo	4.1.1.11	26.19	1.5	27.7

b)

EC Number	Target Species	Target gene	Query gene	Query gene function	EC Number	Seq. similarity	e-value	blast score
2.7.1.45	Gt_Rast							
	G_kaustophilus	GK1957	RTMO02207	## Ribokinase-Gt_GB_Ergo ## Ribokinase (EC 2.7.1.15)-Gt_Embl_Ergo	2.7.1.15	27.88	7e-18	84.3
			RTMO00375	## SpaK-Gt_GB_Ergo ## Subtilin biosynthesis sensor protein spaK (EC 2.7.13.3)-Gt_Embl_Ergo	2.7.13.3	25.93	1.2	27.3
			RTMO01784	## 1-phosphofructokinase-Gt_GB_Ergo ## 1-phosphofructokinase (EC 2.7.1.56)-Gt_Embl_Ergo	2.7.1.56	27.52	1.4	26.9
		GTNG_1768	RTMO02207	## Ribokinase-Gt_GB_Ergo ## Ribokinase (EC 2.7.1.15)-Gt_Embl_Ergo	2.7.1.15	30.18	7e-21	94.4
			RTMO01784	## 1-phosphofructokinase-Gt_GB_Ergo ## 1-phosphofructokinase (EC 2.7.1.56)-Gt_Embl_Ergo	2.7.1.56	23.55	7e-06	44.7
			RTMO00255	## Carbohydrate diacid regulator-Gt_GB_Ergo ## Carbohydrate diacid regulator-Gt_Embl_Ergo		32.61	0.93	27.7
	G_thermodenitrificans	GTNG_1857	RTMO02207	## Ribokinase-Gt_GB_Ergo ## Ribokinase (EC 2.7.1.15)-Gt_Embl_Ergo	2.7.1.15	28.62	4e-15	75.1
			RTMO00582	## Sensory box/GGDEF family protein-Gt_GB_Ergo ## Sensory box/GGDEF family protein-Gt_Embl_Ergo		28.57	0.81	27.7
			RTMO00726	## hypothetical protein-Gt_GB_Ergo ## -Gt_Embl_Ergo		47.83	1.3	26.9
		GTNG_1486	RTMO02207	## Ribokinase-Gt_GB_Ergo ## Ribokinase (EC 2.7.1.15)-Gt_Embl_Ergo	2.7.1.15	26.06	3e-12	65.9
			RTMO01784	## 1-phosphofructokinase-Gt_GB_Ergo ## 1-phosphofructokinase (EC 2.7.1.56)-Gt_Embl_Ergo	2.7.1.56	23.05	0.034	32.3
			RTMO02611	## 3-ketoacyl-CoA thiolase-Gt_GB_Ergo ## 3-ketoacyl-CoA thiolase (EC 2.3.1.16)-Gt_Embl_Ergo	2.3.1.16	34.09	0.15	30.0
	G_WCH70							
	G_Y412MC61	GYMC61_2704	RTMO02207	## Ribokinase-Gt_GB_Ergo ## Ribokinase (EC 2.7.1.15)-Gt_Embl_Ergo	2.7.1.15	28.48	3e-19	88.6
			RTMO01784	## 1-phosphofructokinase-Gt_GB_Ergo ## 1-phosphofructokinase (EC 2.7.1.56)-Gt_Embl_Ergo	2.7.1.56	24.92	8e-08	50.8
			RTMO02818	## Phosphomethylpyrimidine kinase-Gt_GB_Ergo ## Phosphomethylpyrimidine kinase (EC 2.7.4.7) / Hydroxymethylpyrimidine kinase (EC 2.7.1.49)-Gt_Embl_Ergo	2.7.4.7 2.7.1.49	23.03	0.22	29.6
	B_subtilis	BSU22110	RTMO02207	## Ribokinase-Gt_GB_Ergo ## Ribokinase (EC 2.7.1.15)-Gt_Embl_Ergo	2.7.1.15	24.73	3e-06	45.8
			RTMO00480	## Hypothetical cytosolic protein-Gt_GB_Ergo ## Hypothetical cytosolic protein-Gt_Embl_Ergo		32.35	1.1	27.3
			RTMO00284	## Hypothetical protein-Gt_GB_Ergo ## Hypothetical protein-Gt_Embl_Ergo		39.29	1.3	26.9
	E_coli	b3526	RTMO02207	## Ribokinase-Gt_GB_Ergo ## Ribokinase (EC 2.7.1.15)-Gt_Embl_Ergo	2.7.1.15	26.44	3e-07	49.3
			RTMO01784	## 1-phosphofructokinase-Gt_GB_Ergo ## 1-phosphofructokinase (EC 2.7.1.56)-Gt_Embl_Ergo	2.7.1.56	34.92	0.005	35.0
			RTMO03219	## DnaK suppressor protein-Gt_GB_Ergo ## DnaK suppressor protein-Gt_Embl_Ergo		31.82	1.5	26.9

Figure 3.19: Bidirectional BLAST results for altronate dehydratase (a) and 2-dehydro-3-deoxygluconokinase (b) against using corresponding gene sequences from *G_thermoglucosidasius* C56-YS93 (G_thermoglucosidasius), *G. kaustophilus* HTA426 (G_kaustophilus), *G. thermodenitrificans* CCB-US.3-UF5 , *Geobacillus* sp. Y412MC61 (G_Y412MC61), *Geobacillus* sp. WCH70 (G_WCH70), *Bacillus subtilis* 168 (B_subtilis) and *Escherichia coli* (E_coli) .

EC Number	Target Species	Target gene	Query gene	Query gene function	EC Number	Seq. similarity	e-value	blast score
4.1.2.14	Gt_Rast							
	G_kaustophilus	GK1956	RTMO03650	## Thiamin-phosphate pyrophosphorylase-Gt_GB_Ergo ## Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)-Gt_Embl_Ergo	2.5.1.3	31.25	0.013	33.1
			RTMO01874	## Thiamin-phosphate pyrophosphorylase-Gt_GB_Ergo ## Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)-Gt_Embl_Ergo	2.5.1.3	35.56	0.44	28.1
			RTMO01495	## putative nucleoside-diphosphate-sugar epimerases-Gt_GB_Ergo ## putative nucleoside-diphosphate-sugar epimerases-Gt_Embl_Ergo		23.56	0.71	27.3
	G_thermodenitrificans	GTNG_1488	RTMO00399	## Lipoprotein (pheromone precursor)-Gt_GB_Ergo ## Lipoprotein (pheromone precursor)-Gt_Embl_Ergo		24.30	0.007	33.9
			RTMO03650	## Thiamin-phosphate pyrophosphorylase-Gt_GB_Ergo ## Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)-Gt_Embl_Ergo	2.5.1.3	33.73	0.063	30.8
			RTMO01874	## Thiamin-phosphate pyrophosphorylase-Gt_GB_Ergo ## Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)-Gt_Embl_Ergo	2.5.1.3	41.03	0.100	30.0
		GTNG_1767	RTMO01874	## Thiamin-phosphate pyrophosphorylase-Gt_GB_Ergo ## Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)-Gt_Embl_Ergo	2.5.1.3	43.24	0.054	30.8
			RTMO03650	## Thiamin-phosphate pyrophosphorylase-Gt_GB_Ergo ## Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)-Gt_Embl_Ergo	2.5.1.3	33.70	0.078	30.4
			RTMO04040	## Ribulose-phosphate 3-epimerase-Gt_GB_Ergo ## Ribulose-phosphate 3-epimerase (EC 5.1.3.1)-Gt_Embl_Ergo	5.1.3.1	44.44	0.54	27.7
			RTMO03650	## Thiamin-phosphate pyrophosphorylase-Gt_GB_Ergo ## Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)-Gt_Embl_Ergo	2.5.1.3	33.33	0.001	36.2
		GTNG_1856	RTMO01874	## Thiamin-phosphate pyrophosphorylase-Gt_GB_Ergo ## Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)-Gt_Embl_Ergo	2.5.1.3	40.00	0.009	33.5
			RTMO00052	## Hypothetical protein-Gt_GB_Ergo ## Hypothetical protein-Gt_Embl_Ergo		29.27	0.93	26.9
	G_WCH70							
	G_Y412MC61	GYMC61_2703	RTMO01874	## Thiamin-phosphate pyrophosphorylase-Gt_GB_Ergo ## Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)-Gt_Embl_Ergo	2.5.1.3	36.59	0.016	32.7
			RTMO03650	## Thiamin-phosphate pyrophosphorylase-Gt_GB_Ergo ## Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)-Gt_Embl_Ergo	2.5.1.3	26.62	0.15	29.3
			RTMO03608	## DNA segregation ATPase and related proteins (FtsK/SpoIIIE family)-Gt_GB_Ergo ## DNA segregation ATPase and related proteins (FtsK/SpoIIIE family)-Gt_Embl_Ergo		42.11	0.20	28.9
	B_subtilis	BSU22100	RTMO00673	## OppD-Gt_GB_Ergo ## Oligopeptide transport ATP-binding protein oppD-Gt_Embl_Ergo		26.67	0.23	28.5
			RTMO03121	## Low-affinity inorganic phosphate transporter-Gt_GB_Ergo ## Low-affinity inorganic phosphate transporter-Gt_Embl_Ergo		50.00	1.1	26.6
			RTMO04610	## FtsZ-Gt_GB_Ergo ## Cell division protein ftsZ-Gt_Embl_Ergo		25.81	2.4	25.4
			RTMO00611	## Calcium/proton antiporter-Gt_GB_Ergo ## Calcium/proton antiporter-Gt_Embl_Ergo		27.78	0.28	28.5
	E_coli	b1850	RTMO00671	## OppB-Gt_GB_Ergo ## Oligopeptide transport system permease protein oppB-Gt_Embl_Ergo		31.71	0.31	28.5

Figure 3.20: Bidirectional BLAST results for 2-dehydro-3-deoxy-phosphogluconate aldolase (c) against using corresponding gene sequences from *G_thermoglucosidasius* C56-YS93 (G_thermoglucosidasius), *G. kaustophilus* HTA426 (G_kaustophilus), *G. thermodenitrificans* CCB-US.3-UF5 , *Geobacillus* sp. Y412MC61 (G_Y412MC61), *Geobacillus* sp. WCH70 (G_WCH70), *Bacillus subtilis* 168 (B_subtilis) and *Escherichia coli* (E_coli) .

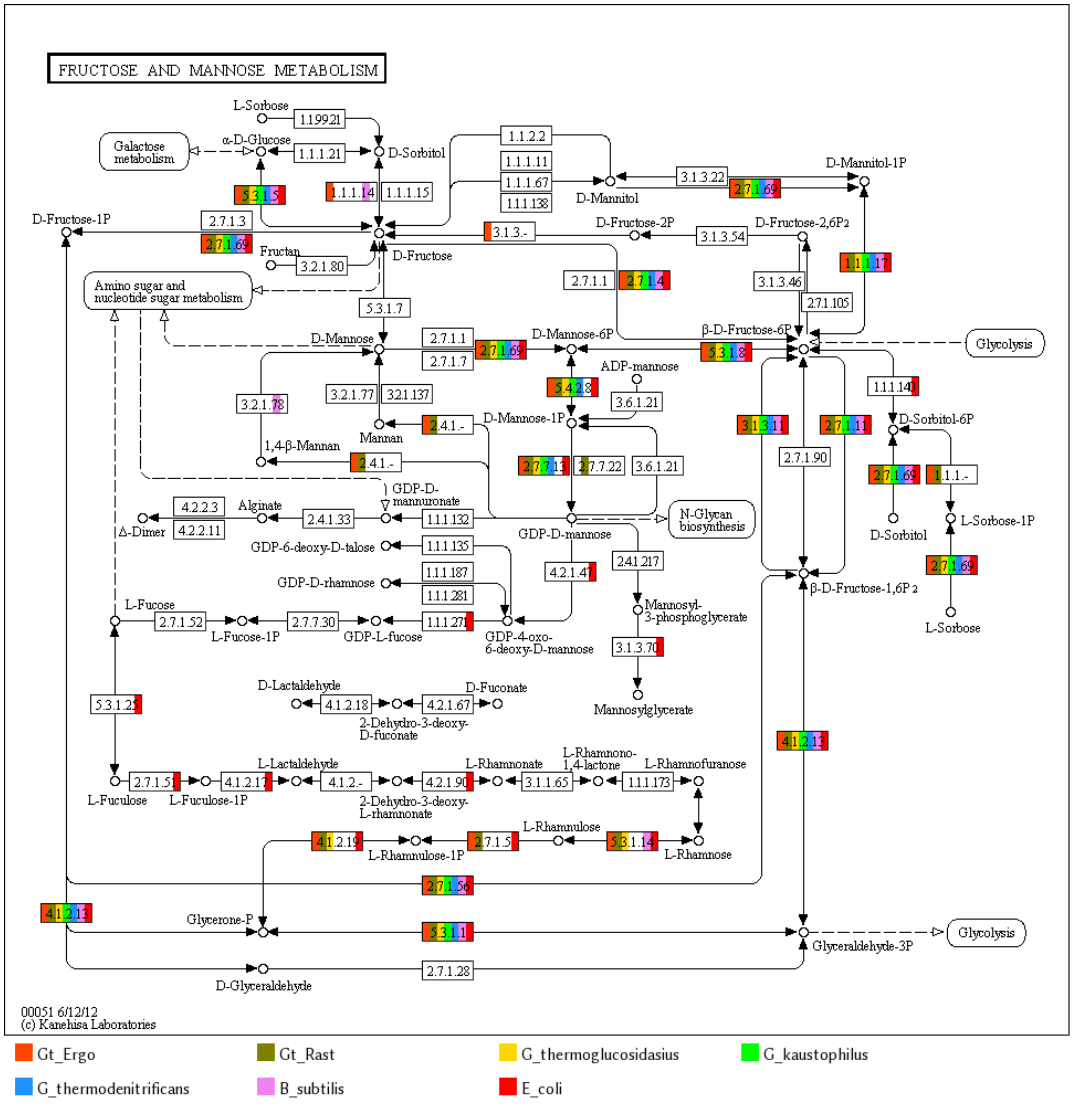


Figure 3.21: PathwayBooster gene annotation for KEGG pathway: Fructose and Mannose Metabolism. Gt_ERGO and Gt_RAST stand for *G. thermoglucosidasius* NCIMB 11955 annotation done by ERGO Integrated Genomics (Gt_ERGO) and RAST SEED server (Gt_RAST) respectively.

3.2.6 Starch and sucrose metabolism

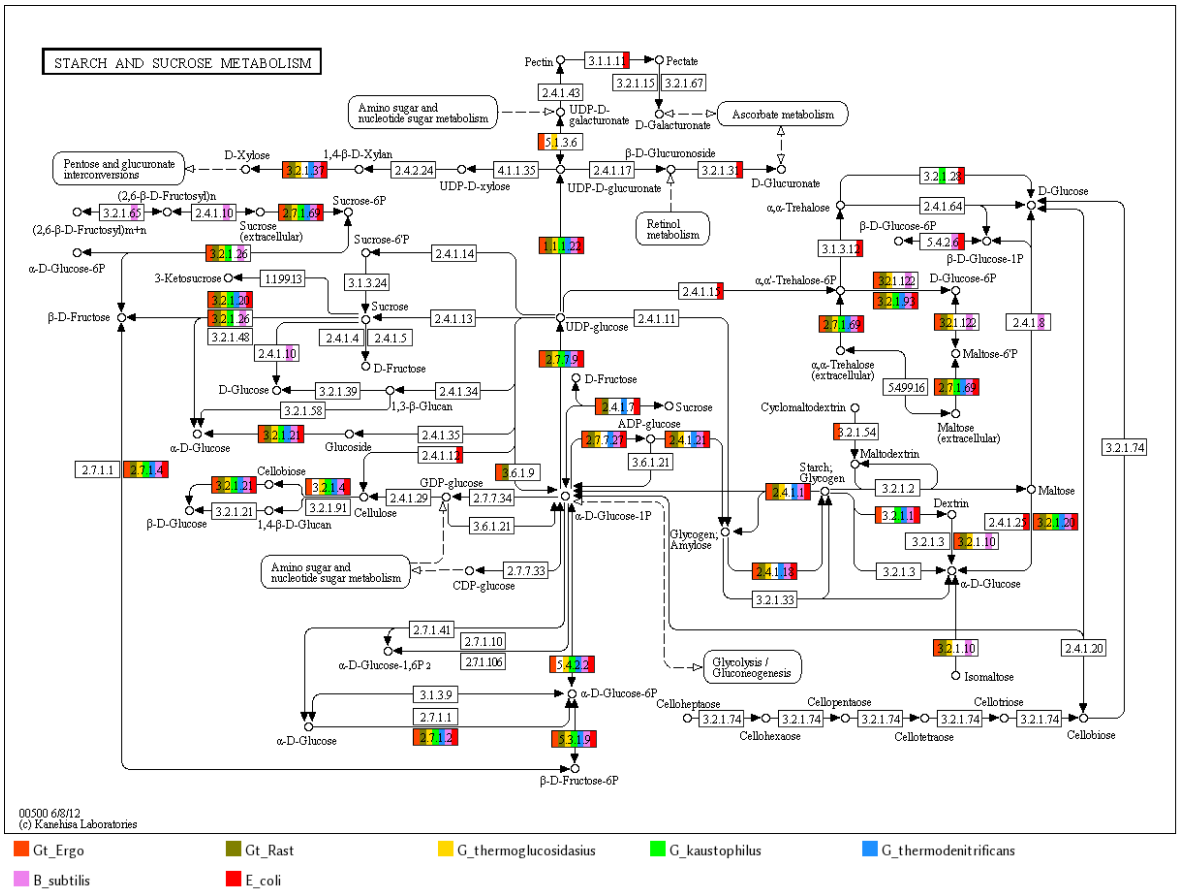


Figure 3.22: PathwayBooster gene annotation for KEGG pathway: Starch and Sucrose Metabolism. Gt_ERGO and Gt_RAST stand for *G. thermoglucosidasius* NCIMB 11955 annotation done by ERGO Integrated Genomics (Gt_ERGO) and RAST SEED server (Gt_RAST) respectively.

G. thermoglucosidasius NCIMB 11955 has annotated nucleotide-sugar pyrophosphatase (EC 3.6.1.9). This enzyme is acting in pathways such as starch and sucrose metabolism, purine metabolism, riboflavin metabolism, nicotinate and nicotinamide metabolism, and pantothenate and CoA biosynthesis. Non canonical nucleoside triphosphate pyrophosphatases release pyrophosphate and monophosphate such as nucleoside triphosphates (NTPs) by hydrolysing the phosphoanhydride bond of corresponding non canonical NTPs such as inosine triphosphate (ITP), deoxyinosine triphosphate (dITP) or xanthosine monophosphate (XMP) [125, 30]. If accumulated, NTPs cause damage to the cell, such as interference with reactions requiring ATP or being undesirably incorporated into DNA and RNA [9]. NTPases are known as the "house-cleaning enzymes" of the cell since they can discriminate between the NTPs and ATP or GTP and hence maintain a healthy pool of DNA and RNA [10]. This

enzyme catalyses an alternative conversion of UDP-glucose to α-D-glucose-1-phosphate in starch and sucrose metabolism pathway through the following general reaction:

$$dinucleotide + H_2O \leftrightarrow 2mononucleotides \tag{3.10}$$

However, it shows high propensity for a range of substrates such as: NAD^+, $NADP^+$, FAD, CoA, ATP and ADP. This enzyme is predicted to facilitates six different reactions within the metabolism of *G. thermoglucosidasius* NCIMB 11955 in four different pathways. For nicotinate and nicotinamide metabolism (reactions 3.11-3.16) in such pathways as riboflavin metabolism, pyrimidine metabolism and purine metabolism. This pyrophosphatase has not been found in any other model organism used in this comparative analysis. It should be noted that this enzyme belongs to a family of noncanonical NTPases that had been found in archaea, bacteria and eukarya [10].

$$NAD^+ + H_2O \leftrightarrow AMP + Nicotinamide - D - ribonucleotide \tag{3.11}$$

$$Deamino - NAD^+ + H_2O \leftrightarrow AMP + Nicotinate - D - ribonucleotide \tag{3.12}$$

$$FAD + H_2O \leftrightarrow AMP + FMN \tag{3.13}$$

$$FAD + H_2O \leftrightarrow AMP + FMN \tag{3.14}$$

$$UDP - glucose + H_2O \leftrightarrow UMP + D - Glucose - 1 - phosphate \tag{3.15}$$

$$3' - Phosphoadenylylsulfate + H_2O \leftrightarrow Sulfate + Adenosine - 3', 5' - bisphosphate \tag{3.16}$$

G. thermoglucosidasius NCIMB 11955 and *G. kaustophilus* HTA426 have been found to be capable of also catalysing the conversion of sucrose to β-D-glucose-1-phosphate which is facilitated by sucrose phosphorylase (EC 2.4.1.7).

Although all the species used in the comparison can uptake α,α-trehalose and maltose from the extracellular environment, only *G. thermoglucosidasius* strains C56-YS93 and 11955 as well as *Bacillus subtilis* can further break down maltose 6-phosphate to D-glucose-6-phosphate using phospho-α-glucosidase (EC 3.2.1.122). Phospho-α-glucosidase can also facilitate conversion of α,α'-trehalose 6-phosphate to D-glucose-6-phosphate (Figure 3.22).

G. thermoglucosidasius NCIMB 11955 can also metabolise isomaltose using oligo-1,6-glucosidase (EC 3.2.1.10). This enzyme also allows the conversion of dextrin to α-D-glucose, which is the second step in conversion of starch to glycogen. (Figure 3.22).

G. thermoglucosidasius NCIMB 11955 along with some other strains have α-amylase and neopullulanase genes positioned close to one another on the genome. Both of these enzymes play a crucial role in starch metabolism with the neopullulanase being responsible for the hydrolysis of 1-6 linked side chains [60, 59]. G. thermoglucosidasius NCIMB 11955 can also take up extracellular isomaltose using oligo- 1,6-glucosidase (EC 3.2.1.10). This enzyme catalyses the conversion from dextrin to α-D-glucose, which is a second step in the conversion of starch to glycogen [60, 59].

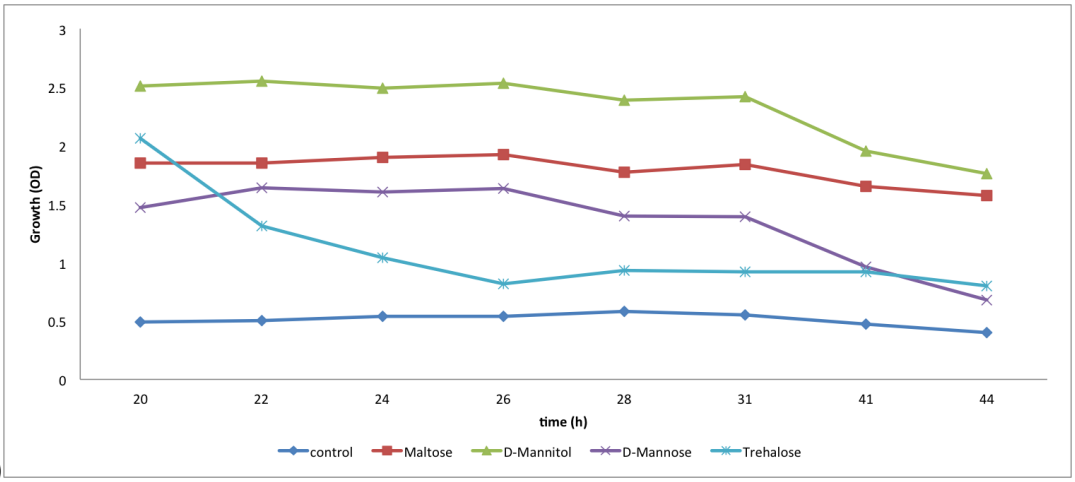
It is only within the two substrains of *G. thermoglucosidasius* that a conversion from UDP-D-galacturonate to UDP-D-glucoronate is possible using UDP-glucuronate 4-epimerase (EC 5.1.3.6) within starch and sucrose metabolism pathway as well as in amino sugar and nucleotide sugar metabolism.

Theoretical predictions of carbohydrate metabolism of *G. thermoglucosidasius* NCIMB 11955 has been confirmed experimentally through strain growth on different carbohydrates (Figure 3.23). This bacterial strain was found to be able to metabolise as discussed in the previous sections: L-rhamnose, D-maltose and D-thehalose. The experimental growth of the strain on different carbohydrates showed a strong positive results on 21 out of 49 carbohydrates tested. Four of the carbohydrates resulted in a weak positive growth of *G. thermoglucosidasius* NCIMB 11955, namely: D-melibiose, D-tagatose, potassium 2-ketogluconate and D-galactose. The weak positive test for these carbohydrates is surprising since the bacterium seem to have complete pathways for breakdown of D-melibiose, D-tagatose and D-galactose in its annotated genomes. 16 of the tested carbohydrates resulted in no growth of *G. thermoglucosidasius* NCIMB 11955. These are not surprising since *G. thermoglucosidasius* lacks the gene annotations that would allow for a breakdown of such carbohydrates as D-arabinose, D-adonitol or inositol. Especially since the the genes responsible for the breakdown of the arabinose are located in the HUS region, which in clade "thermoglucosidasius" underwent a significant rearrangements and as a result lost the cluster responsible for the arabinose degradation

	GLY	ERY	DARA	LARA	RIB	DXYL	LXYL
11955	P			P	P	P	
TM242	V			P	P	P	
	ADO	MDX	GAL	GLU	FRU	MNE	SBE
11955			V	P	P	P	
TM242				P	P	P	
	RHA	DUL	INO	MAN	SOR	MDM	MDG
11955	P			P	P		P
TM242				P	P		P
	NAG	AMY	ARB	ESC	SAL	CEL	MAL
11955	P	P	P	P	P	P	P
TM242	P	P	P	P	P	P	P
	LAC	MEL	SAC	TRE	INU	MLZ	RAF
11955		V	P	P			
TM242			P	P			
	AMD	GLYG	XLT	GEN	TUR	LYX	TAG
11955	P			P	P		V
TM242	P			P	P		
	DFUC	LFUC	DARL	LARL	GNT	2KG	5KG
11955						V	P
TM242							

Figure 3.23: Carbohydrate metabolism range of *G. thermoglucosidasius* NCIMB 11955. "P" denotes strong positive result and "V" stands for a weak positive result. The following abbreviations denote the following: GLY for glycerol, ERY for erythritol, DARA for D-arabinose, LARA for L-arabinose, RIB for D-ribose, DXYL for D-xylose, LXYL for L-xylose, ADO for D-adonitol, MDX for methyl-βD-xylopyranose, GAL for D-galactose, GLU for glucose, FRU for fructose, MNE for D-mannose, SBE for L-sorbose, RHA for L-rhamnose, DUL for dulcitol, INO for inositol, MAN for D-mannol, SOR for sorbitol, MDM for methyl-αD-mannopyranoside, MDG for methyl-αD-glucopyranoside, NAG for N-acetylglucosamine, AMY for amygdalin, ARB for arbutin, ESC for ferric acid, SAL for salicin, CEL for D-cellobiose, MAL for D-maltose, LAC for D-lactose, MEL for D-melibiose, SAC for D-saccharose, TRE for D-trehalose, INU for inulin, MLZ for D-melezitose, RAF for D-Raffinose, AMD for starch, GLYG for glycogen, XLT for xylitol, GEN for gentiobiose, TUR for D-turanose, LYX for D-lyxose, TAG for D-tagatose, DFUC for D-fucose, LFUC for L-fucose, DARL for D-arabitol, LARL for L-arabitol, GNT for potassium gluconate, 2KG for potassium 2-ketogluconate and 5KG for potassium 5-ketogluconate. TM242 is a *G. thermoglucosidasius* NCIMB 11955 mutant strain with lactate dehydrogenase knock-out [38].

a)



b)

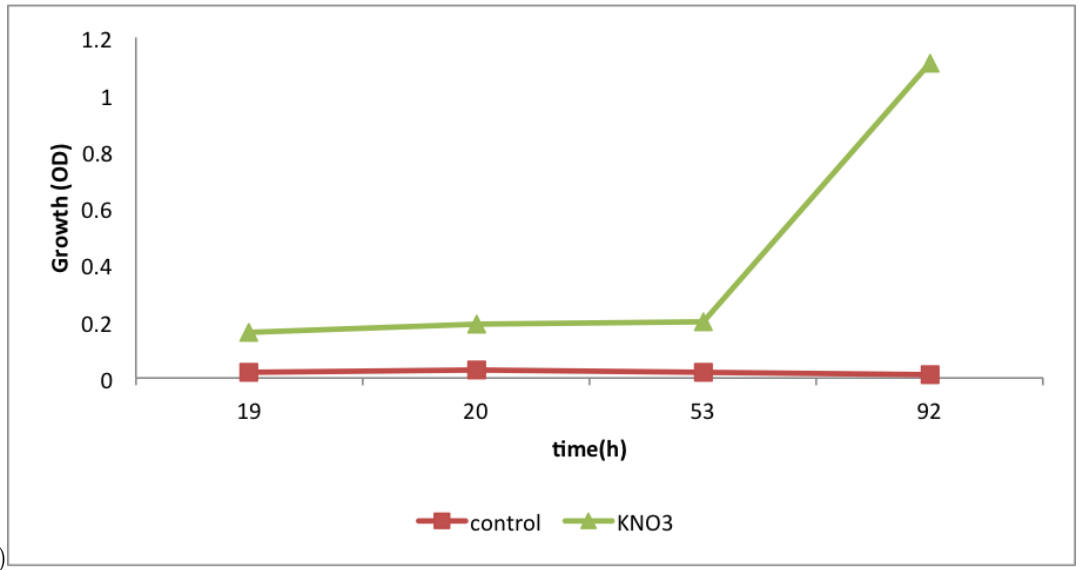


Figure 3.24: D-mannitol, maltose, D-mannose and trehalose growth (a) and nitrate(b) assays results.The controls show the cultures which were grown without the addition of nitrate(a) or maltose, D-mannitol, d-mannose or trehalose.

3.2.7 Other unique metabolic assignments

Other notable enzymes which give *Geobacillus thermoglucosidasius* 11955 a unique metabolic profile are: anaerobic carbon monoxide dehydrogenase (EC 1.2.99.2) which is shared by *G. kaustophilus*, nitrilase (EC 3.5.5.1) with is only present in *G. thermoglucosidasius* strains 11955 and C56-YS93 and nicotinamidase (EC 3.5.1.19) that is found only in strain 11955 and *E.coli* (when compared to the strains used for the PathwayBooster comparative analysis).

In pyrimidine metabolism strain 11955 displays unique annotations, shared with no other members of neither *Geobacillus* nor *Bacillus*. These include dihydropyrimidine dehydrogenase ($NADP^+$) (EC 1.3.1.2), pseudouridylate synthase (EC 4.2.1.70) and beta-ureidopropionase (EC 3.5.1.6). Similarly, methylenetetrahydrofolate reductase [NAD(P)H] (EC 1.5.1.20) which is found in a carbon fixation pathways in prokaryotes can be found only in strain 11955. In the same pathway formyltetrahydrofolate synthetase (EC 6.3.4.3) is found assigned uniquely to *G.thermoglucosidasius* strains.

Caution was exercised when dealing with annotations that cannot be found within closely related species as they might indicate error in the gene assignment. However, when both software algorithms are in agreement on a given gene annotation and a given metabolic route does not contain a missing step that could be otherwise filled by bidirectional BLAST analysis, then the annotation is considered as a viable possibility.

3.3 Vitamin B_{12}biosynthesis

3.3.1 Canonical routes for *de novo* vitamin B_{12}biosynthesis

Cobalamin is a structurally-complex cofactor with a tetrapyrrole-derived framework and a chelated cobalt ion at its core. The structure is based on a corrin, which serves as a skeleton for this compound where the cobalt atom is surrounded on one side by four nitrogen atoms forming a tetrapyrolle ring and on the other by nitrogen derived from benzimidazole. The most common form in which vitamin B_{12} is extracted naturally is cyanocobalamin, where in the sixth coordination position the substituent is -CN. This substituent can be in turn replaced by -OH, -Cl, -NO_2 or -CNS.

Adenosyl-cobalamin (vitamin B_{12} form) is an essential cofactor that plays a major role in a number of core enzymes [107]. These cobalamin-dependent enzymes in *G. thermoglucosidasius* NCIMB 11955 include, amongst others: methylmalonyl-CoA mutase, methionine synthase (cobalamin-dependent), propanediol dehydratase or methylmalonyl-CoA mutase. Even though the majority of bacteria use enzymes that require cobalamin, only a few can synthesise it *de novo* [107].

In summary, the steps responsible for conversion from uroporphyrinogen III to cobyric acid involve at least eight C-methylations, a C-20 elimination, an NADPH-dependent macrocycle reduction, a C-12 acetate decarboxylation, the migration of methyl from C-11 to C-12 and cobalt insertion([34]).

Two major routes of cobalamin biosynthesis have been described in the literature, which carry out these steps in a different order and they are divided into aerobic (oxygen-dependent) and anaerobic (oxygen-independent) routes [107]. The point of formation of the corrin ring and the cobalt incorporation are the most crucial differences between the routes. In the aerobic path the corrin ring synthesis occurs before cobalt insertion. Conversely, in the anaerobic pathway this step is reversed: cobalt is inserted into precorrin-2 as one of the first steps in conversion from uroporphyrinogen to cobyric acid [107, 34]. The next steps in the cobalamin synthesis involve adenosylation of the corrin ring, attachment of an aminopropanal arm, and assembly of a nucleotide loop that bridges dimethylbenzimidazole and the corrin ring ([107]), all of which are shared between the two routes. The genes that are usually associated with anaerobic routes are referred to as "Cbi" genes whilst "Cob" genes represent the aerobic version of a pathway. These genes are found usually within highly conserved operons. In the literature a nonspecific transporter CorA has been described as responsible for uptake of

cobalt. The two canonical routes for synthesis of cobalamin are shown in Figure 3.25.

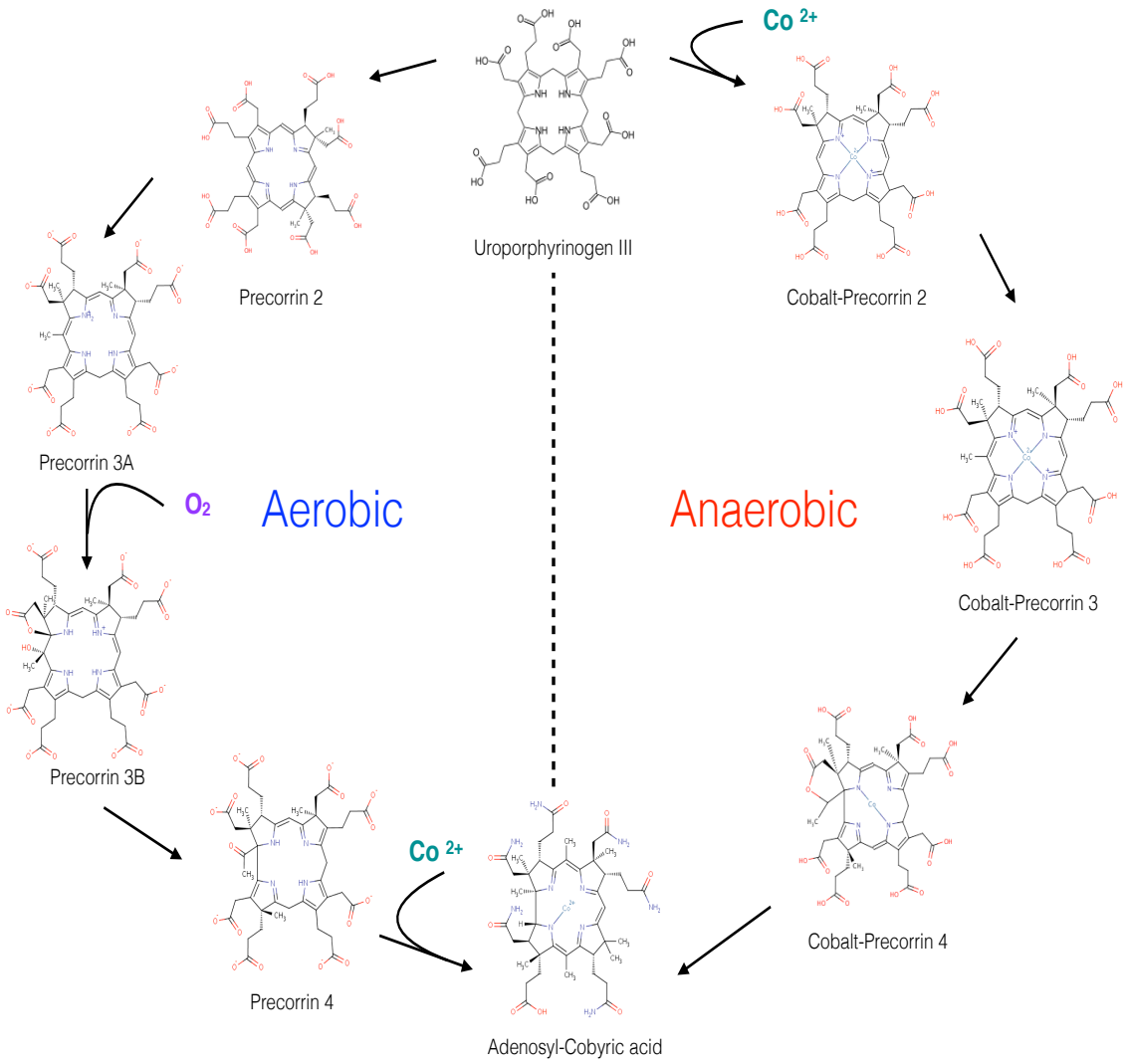


Figure 3.25: Canonical aerobic (left-hand side) and anaerobic (right-hand side) biosynthesis routes for biosynthesis of cobyric acid from uroporpyrinogen III.

3.3.2 Proposed unconventional vitamin B_{12} biosynthesis route in the genus *Geobacillus*

It has been observed that under laboratory conditions *G. thermoglucosidasius* NCIMB 11955 does not need the addition of vitamin B_{12} in the growth medium. In the light of enzymatic capabilities of this bacterium and the presence of crucial cobalamin-dependent enzymes, it was evident that the strain must be able to synthesise *de novo* adenosyl-cobalamin.

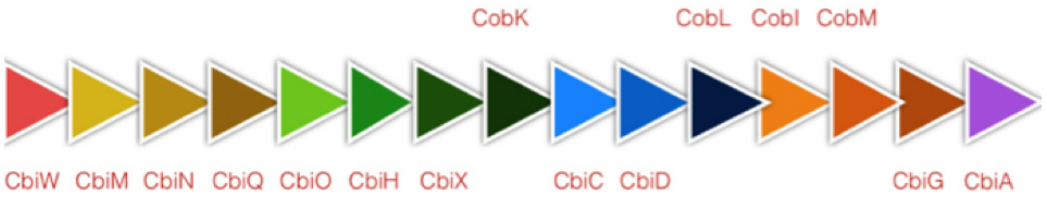


Figure 3.26: Vitamin B_{12} gene operon found within the genome of *G. thermoglucosidasius* NCIMB 11955.

Upon investigation of the gene operon of *G. thermoglucosidasius* NCIMB 11955 for vitamin B_{12} biosynthesis, it was observed that a combination of "Cob" and "Cbi" genes were present (see Figure 3.26). A widened search of this operon within other members of the genus "Geobacillus" revealed that such a combinatorial gene presence was found in all the members. These include strains that are facultative anaerobes and obligate aerobes alike (see Figure (3.27, 3.29).

The genes *cobK*, *cobL*, *cobI* and *cobM* are annotated as precorrin-6x reductase (*cobK*), *cobL* is responsible for conversion of preccorrin-6 to precorrin-8, precorrin-2 C20 methyltransferase (cobI) and precorrin-4 C11-methyltransferase (*cobM*). The closest relative to *G. thermoglucosidasius* NCIMB 11955 with well characterised cobalamin biosythesis is *Bacillus megaterium*; however, this strain has been found to have annotated only the "Cbi" gene operon rather than mixed gene operon described in here. The proposed pathway that would accommodate such a mixed gene operon is presented in Figure 3.28.

The critical step in the proposed mixed pathway is the incorporation of cobalt into precorrin 5, which can be achieved by using sirohydrochlorin cobaltochelatase (EC 4.99.1.3). It has been found to be used by *Salmonella* sp. as a way to incorporate cobalt into pretorian 2 to yield cobalt-precorrin 2 [106].

This is an unconventional route to biosynthesis of vitamin B_{12} and as the Figure 3.29 shows it is highly conserved within the genus. This gene operon is located in one of the three regions

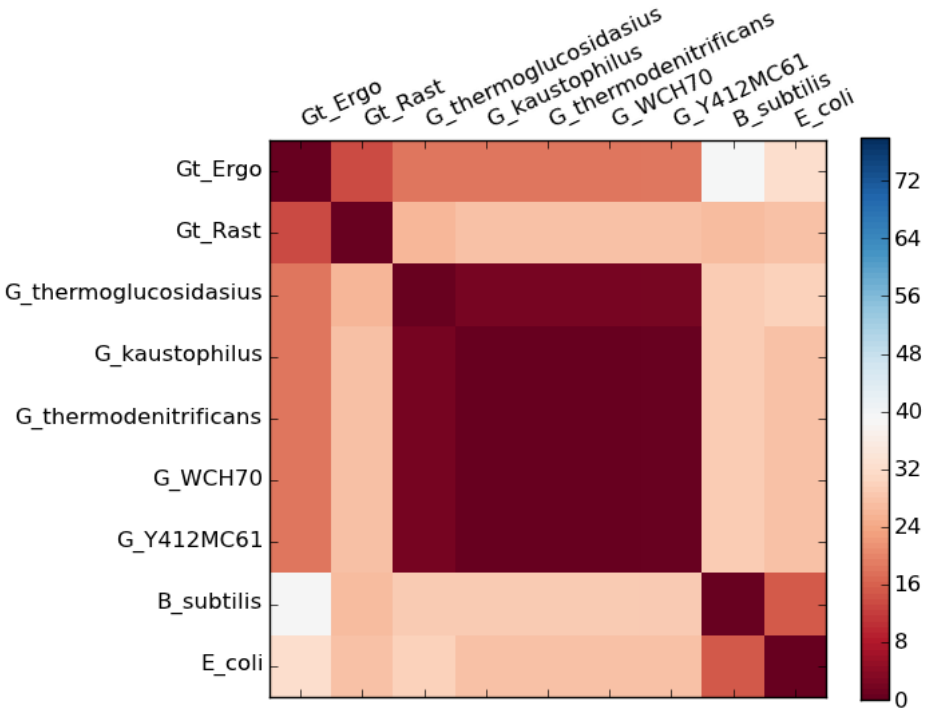


Figure 3.27: Heat map showing gene conservation (based on gene assignment for this route) amongst selected members within genus *Geobacillus* along with *B. subtilis* and *E.coli* for vitamin B_{12}biosynthesis. The scale shows the difference in the gene assignment (the higher the score, the lower the similarity).

of the genome that is unaffected by the genome rearrangement across the *Geobacillus*. In the literature only two routes lead to the biosynthesis of vitamin B_{12} and mixed routes like the one proposed here, has not been described before. This predicted route requires addition research to established what the origins of the anaerobic "Cob" genes are. Furthermore, by knocking-out sirohydrochlorin cobaltochelatase, it could be established if this gene is indeed responsible for the incorporating the cobalt atom into the corrin ring or if this step is catalysed by other enzyme. Keeping these questions in mind, this is a truly novel finding that challenges the conventional understanding of vitamin B_{12} biosynthesis.

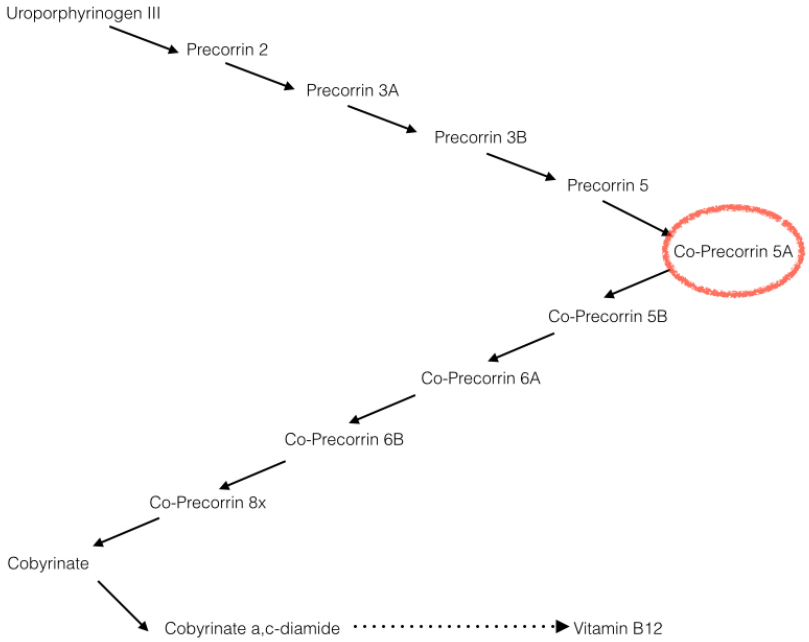


Figure 3.28: Proposed mixed pathway for cobyric acid biosynthesis in *G. thermoglucosidasius* NCIMB 11955

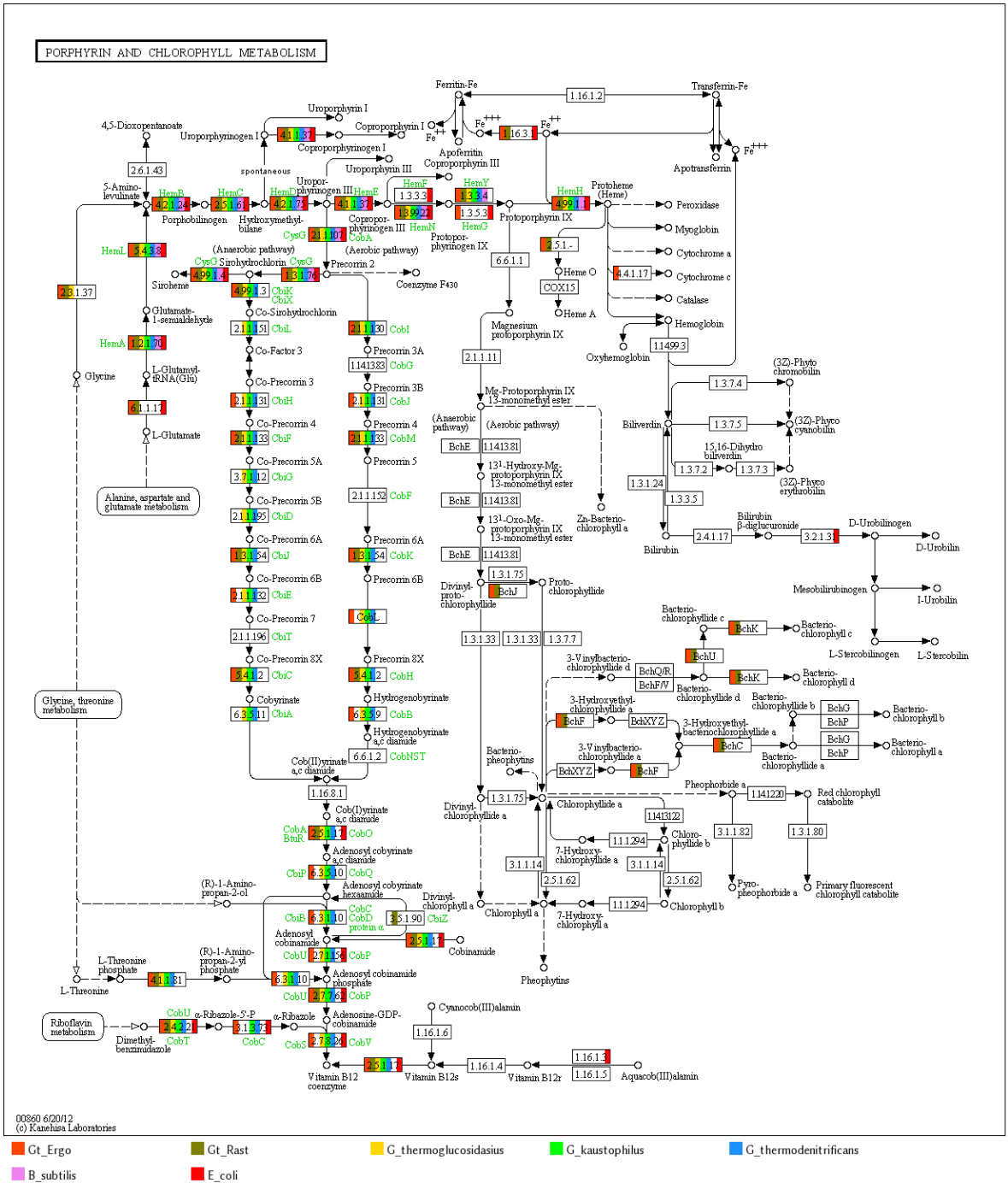


Figure 3.29: Gene assignment for vitamin B_{12} biosynthesis within genus *Geobacillus*.

3.4 Bacterial microcompartments: propanediol utilization operon.

Bacterial microcompartments are polyhedral structures (when viewed in the electron microscope) that encapsulate enzymes in a thin shell [39]. These microcompartments usually contain enzymes that can carry enzymatic routes such as degradation of propandiol. In the genome of *Geobacillus thermoglucosidasius* NCIMB 11955, genomic analysis has revealed the presence of a propanediol utilisation ("Pdu") operon with associated protein shell genes such as *pduA*. The protein shells were identified in *Geobacillus thermoglucosidasius* NCIMB 11955 by Dr Steven Bowden (Figure 3.30). It has been reported in the literature that the utilisation of propandiol requires encapsulation due to the toxicity of the breakdown products and subsequent DNA damage cause by propionaldehyde[39, 109]. The studies on *Salmonella enterica serovar Typhimurium* LT2 suggests that the mutant cells (with protein shell genes knocked-out), when grown on propandiol lead to arrested growth of the bacterium [15]. The

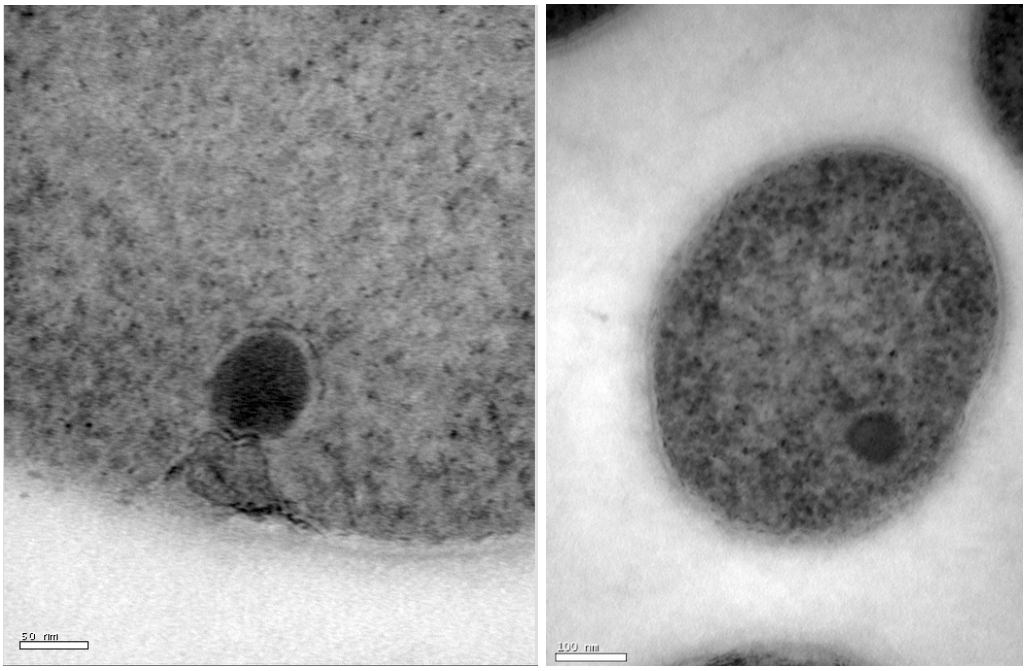


Figure 3.30: Electron microscope image of *Geobacillus thermoglucosidasius* NCIMB 11955 microcompartments. Courtesy of Dr Steven Bowden and Carolyn Williamson.

comparison of the gene operon involved in propanediol utilisation between closely related species within the genus indicate that this operon and compartmentalisation of Pdu pathway can be found only within selected species in the clade "thermoglucosidasius", namely *G.thermoglucosidasius* C56-YS93, NCIM11955 and *Geobacillus sp* Y41MC1 (Figure 3.31).

The gene names and their predicted function are presented in Table 3.2

A BLAST (nucleotide BLAST and BLASTx) analysis suggests that the operon bears close resemblance to that from *Listeria* spp. (see Figure 3.31) [145]. The operon lacks the annotation of *pduW*, which encodes for propionate kinase and *pduF*, which facilitates the diffusion of 1,2-propanediol. The function of propionate kinase can, however, be carried out by acetate kinase, which has been found to catalyse the conversion of propionate to propanoyl phosphate, albeit at a decreased rate [15]. The operon is also missing *pduF*, which encodes 1,2-propanediol diffusion facilitator [15].

The observation of microcompartments under electron microscope confirms the annotation of the genes encoding the protein shell, however it does not confirm the functional propanediol utilisation pathway. Additional research would be required to confirm the ability to efficiently degrade propanediol.

Component name	Function	Gene
pduF	1,2-propanediol diffusion facilitator	missing
pocR	Pdu operon transcription activator	Gthg01747, Gthg01746
PduA	Polyhedral body formation	Gthg01745
PduB	Polyhedral body formation	Gthg01744
PduC	Propanediol dehydratase large subunit	Gthg01743
PduD	Propanediol dehydratase medium subunit	Gthg01742
PduE	Propanediol dehydratase small subunit	Gthg01741
PduG	Propanediol dehydratase reactivation factor large subunit	Gthg01740
PduH	Propanediol dehydratase reactivation factor small subunit	Gthg01739
PduK	Polyhedral body formation	Gthg01738
PduJ	Polyhedral body formation	Gthg01737
PduL	Phosphotransacylase	Gthg01736
EutJ	Ethanolamine utilization protein	Gthg01735
PduM	Polyhedral body formation	Gthg01734
PduN	Polyhedral body formation	Gthg01733
PduO	Cob(I)yrinic acid a,c-diamide adenosyltransferase (CobA)	Gthg01732
PduQ	Propanol dehydrogenase	Gthg01729
PduP	CoA-dependent propionaldehyde dehydrogenase	Gthg01731
PduW	Propionate kinase	missing
PduX	L-Theronine kinase	missing

Table 3.2: Protein names with function as encoded in the Pdu gene operon in *Geobacillus thermoglucosidasius* NCIMB 11955

a)

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics

Description	Max score	Total score	Query cover	E value	Ident	Accession
MULTISPECIES: propanediol utilization protein PduB [Geobacillus]	459	459	97%	1e-160	100%	WP_003250381.1
propanediol utilization protein PduB [Bacillus massiliosenegalensis]	389	389	97%	4e-133	86%	WP_019152643.1
propanediol utilization protein PduB [Bacillus azotoformans]	387	387	97%	2e-132	86%	WP_003332112.1
propanediol utilization protein [Bacillus azotoformans]	387	387	97%	3e-132	85%	WP_035197535.1
propanediol utilization protein [Bacillus thermotolerans]	360	360	97%	1e-121	82%	WP_039235380.1
propanediol utilization protein [Clostridiisalibacter paucivorans]	344	344	97%	2e-115	74%	WP_026896001.1
propanediol utilization protein [Clostridiaceae bacterium BRH_c20a]	339	339	96%	3e-113	73%	WP_045660515.1
propanediol utilization protein [Anaerosalibacter sp. ND1]	328	328	97%	4e-109	73%	WP_042682931.1
propanediol utilization protein [Clostridiales bacterium DRI-13]	325	325	96%	6e-108	71%	WP_034420515.1
MULTISPECIES: propanediol utilization protein PduB [Thermoanaerobacter]	317	317	97%	9e-105	72%	WP_003870136.1

b)

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Geobacillus sp. Y4.1MC1, complete genome	1440	1440	100%	0.0	100%	CP002293.1
Geobacillus thermoglucosidasius C56-YS93, complete genome	1434	1434	100%	0.0	99%	CP002835.1
Pelosinus fermentans JBW45, complete genome	358	358	81%	1e-94	73%	CP010978.1
Listeria innocua Clip11262 complete genome, segment 5/12	345	345	80%	6e-91	72%	AL596167.1
Pelosinus sp. UFO1, complete genome	333	333	80%	4e-87	72%	CP008852.1
Listeria monocytogenes strain L2074, complete genome	331	331	80%	1e-86	72%	CP007689.1
Listeria monocytogenes strain CFSAN007956, complete genome	331	331	80%	1e-86	72%	CP011397.1
Listeria monocytogenes J0161, complete genome	331	331	80%	1e-86	72%	CP002001.1
Listeria monocytogenes strain C1-387, complete genome	329	329	80%	5e-86	72%	CP006591.1
Listeria monocytogenes Finland 1998, complete genome	329	329	80%	5e-86	72%	CP002004.1
Listeria monocytogenes strain L1846, complete genome	325	325	80%	6e-85	71%	CP007688.1
Listeria monocytogenes strain L2625, complete genome	325	325	80%	6e-85	71%	CP007687.1
Listeria monocytogenes strain L2676, complete genome	325	325	80%	6e-85	71%	CP007685.1
Listeria monocytogenes strain L2626, complete genome	325	325	80%	6e-85	71%	CP007684.1

Figure 3.31: Nucleotide BLAST (a) search and a BLASTx (b) search on shell protein gene *pduA* and *pduB*.

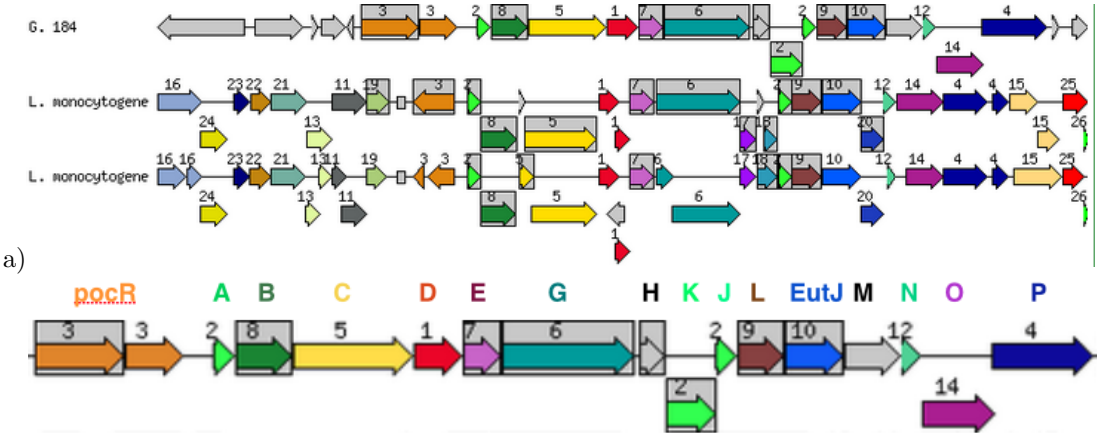
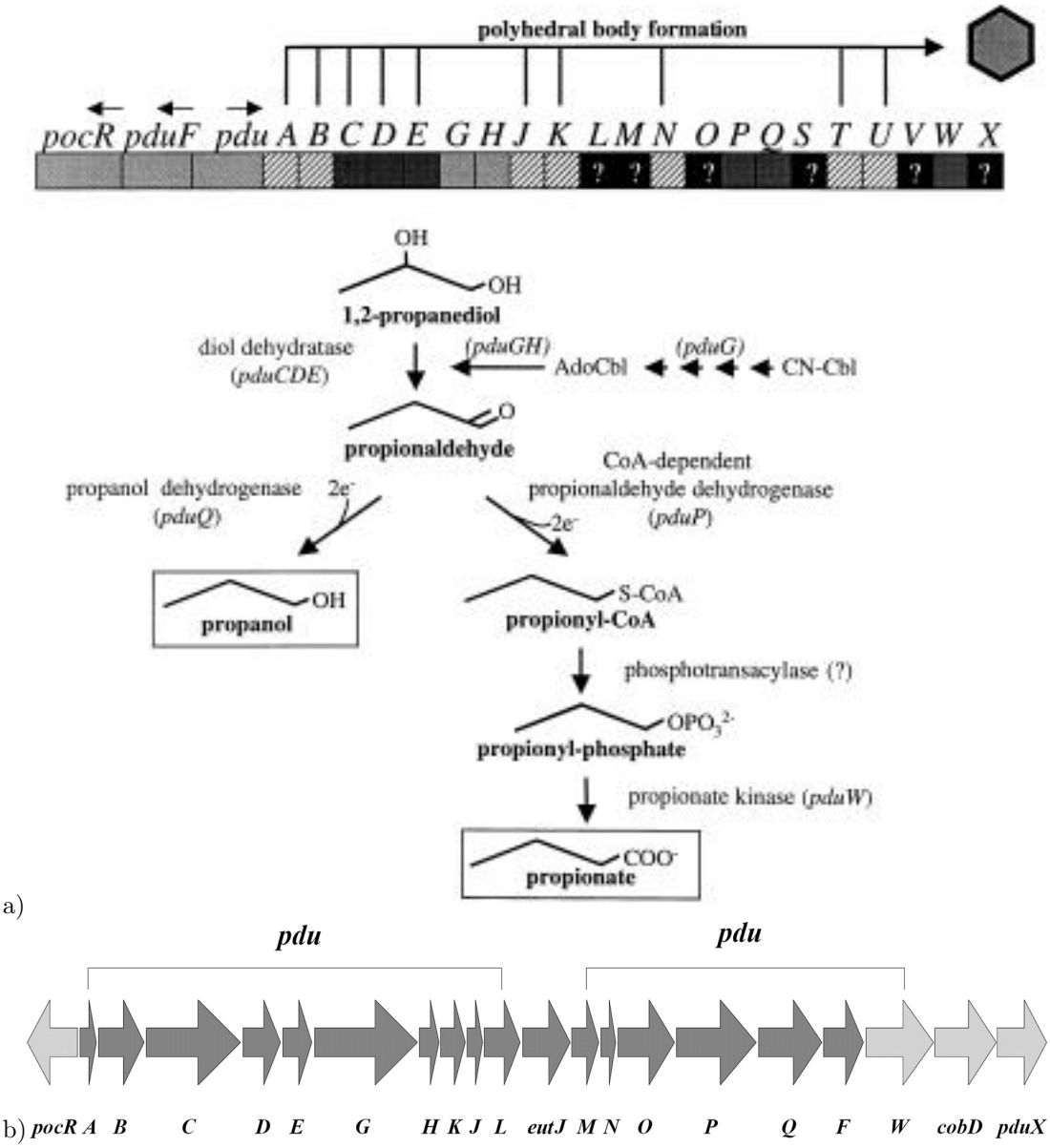


Figure 3.32: 1,2-propanediol utilisation operon. The cluster is compared to those of two strains of *Listeria monocytogene*. The gene annotations (b) are as follows: pocR, pduA, pduB, pduC, pduD, pduE, pduG, pduH, pduJ, pduL, eutJ, pduM, pduN, pduO, pduP. PduQ can be found upstream of the operon before a transposase. The protein functions can be found in Table 3.2



a)

b) *pocR A B C D E G H K J L eutJ M N O P Q F W cobD pduX*

Figure 3.33: 1,2-propanediol utilisation gene operon in *Salmonella enterica* with the functions of the protein products. (a). As reported by Bobik *et al.* 1999 [16] and (b) gene operon as reported in *Listeria innocua*, Xue *et al.* 2008 [145].

3.5 Conclusions

Geobacillus thermoglucosidasius NCIMB 11955 is an interesting bacterium with applications in industry. Although few other *Geobacillus* spp. have been described in the literature, it is novel to look at a comprehensive metabolic comparative analysis and see how different each strain really is. The genus *Geobacillus* displays a few interesting metabolic differences to those found within its mesophilic counterpart. Interestingly, *Geobacillus* spp. have acquired 2-oxoglutarate synthase (EC 1.2.7.3), which can be used in reductive citric acid cycle, but they have also lost the NAD(P)$^+$ transhydrogenase (1.6.1.1 or 1.6.1.2),which makes the conversion from NAD$^+$ to NADP$^+$ and vice versa difficult. This analysis has shown that *Geobacillus* spp have three major parts of genome conservation (as discussed in the previous chapter), where the core metabolic capabilities are shared across the genus. Although such modular genome conservation across species within the same genus is not uncommon [80], in this genus these conserved regions encode enzymes involved in central carbon metabolism, unconventional vitamin B_{12} biosynthesis pathway, core sugar metabolism and enzymes involved in cell maintenance.

Even though a high degree of conservation can be found in those three genome segments, *Geobacillus* spp. exhibit a high degree of difference even on sub-strain level, along with unique metabolic advantages such as pyruvate-formate lyase, GAPDH (EC 1.2.1.59) in *G. thermoglucosidasius* NCIMB 11955 or pyruvate synthase and GAPDH (EC 1.2.1.12) in *G. thermodenitrificans* CCB-US.3-UF5. A distinction can be made between strains belonging to clades associated with aerobic respiration and those capable of growing when oxygen is scarce.

In the context of industrial applications, the metabolic predictions suggest that *G. thermoglucosidasius* NCIMB 11955 can be used commercially to degrade starch. The metabolic predications suggests that this bacterium has α-amylase to break α1-4 linkage. However a complete breakdown of starch requires the breakdown of 1-6 linkages that can be done using thermostable *G. thermoglucosidasius* NCIMB 11955 neopullulanase [60]. Thermostable neopullulanase has the ability to break both α1-4 and 1-6 linkages making it an ideal bacteria for breakdown of starch in one organism, under high temperatures and on industrial scale.

The predicted metabolism however indicates that *G. thermoglucosidasius* NCIMB 11955 could not be use for hemicellulose degradation in industry due to incomplete clusters of HUS locus. Although other strains belonging to obligate aerobes clades have a complete set

of genes to degrade hemicellulose completely, the strains from clade "thermoglucosidasius" could not carry out the first step of degradation.

The ability to degrade catechol, through *G. thermoglucosidasius* NCIMB 11955 plasmid-chromosome pathway has also an interesting application in the industry. Catechol is widely used in industry as a photographic developer, polymerisation inhibitor or lubricating oil and as such needs to be removed from wastewater through various method [2]. *G. thermoglucosidasius* NCIMB 11955 biologically degrade catechol in a safe and efficient manner [2]. Overall the metabolic predictions display a battery of thermostable metabolic enzymes that can be useful in industrial applications.

Chapter 4

Reconstruction of Genome Scale Metabolic Model of *G. thermoglucosidasius* NCIMB 11955

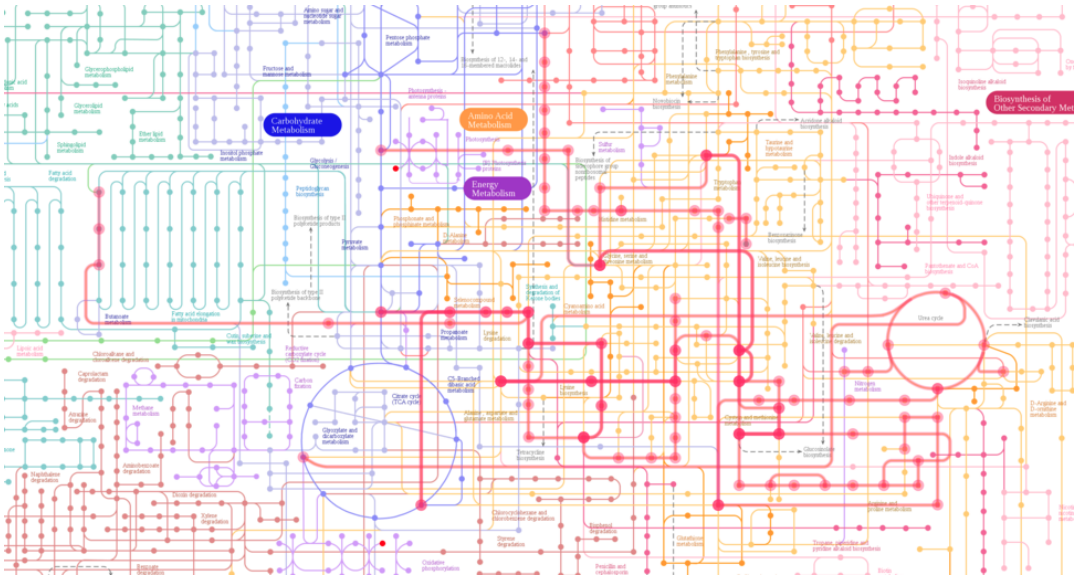


Figure 4.1: Representation of global metabolism of *G. thermoglucosidasius* NCIMB 11955.

4.1 Introduction

Metabolic models were introduced over the last decade and ever since have generated a lot of interest from the scientific community. A genome-scale view of the metabolism, rather than from a singular pathway point of view has become a useful tool. Genome-scale metabolic models (GEMs) can address questions as much about a specific metabolic pathway or single substrate as on a macro scale, when physiological observations need explanations. In this chapter, the reconstruction of a metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 is described, with respect to the findings from the analyses on gene assignments (discussed in the previous chapter). A bottom-up approach was used for the reconstruction,which means building a model from genome annotation. This chapter aims to introduce metabolic model and present initial findings from flux balance analysis (FBA), with respect to unique gene assignments as highlighted in the previous chapter (such as GAPDH or 2-oxoglutarate synthase). This reconstruction elaborates on model refinement techniques within the available *in silico* environments and explains the steps undertaken for the curation of the model. The model reconstruction of *Geobacillus thermoglucosidasius* NCIMB 11955 was based on experimental data for estimation of biomass, lipid composition and biosynthetic costs.

4.2 Reconstruction of metabolic model of *G. thermoglucosidasius* NCIMB 11955

As mentioned previously, the metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 was generated based on a bottom-up approach. This method entails using annotated genome sequence as the basis of the reconstruction. It can be argued that the metabolic model can only be as good as its annotation which is why, in this study a genome annotation comparison was used [134]. Methodology for that comparison and generation of confidence scores is discussed in depth in "Materials and Methods" as well as in the previous chapter. The breakdown of confidence scores can be found in the Figure 4.2

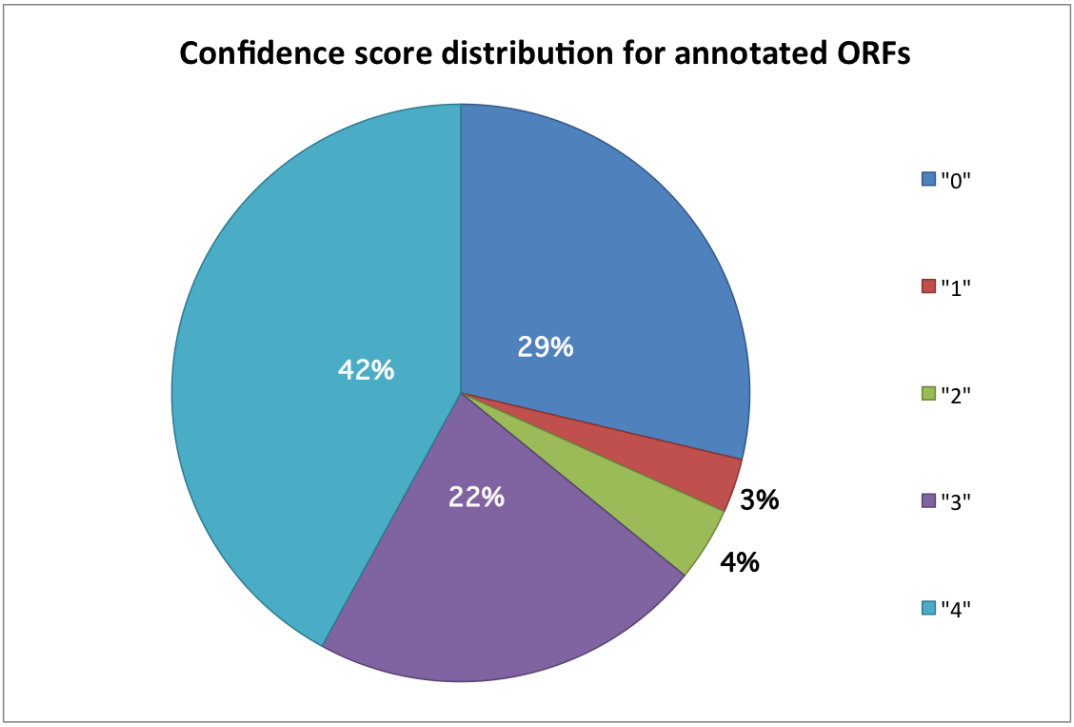


Figure 4.2: Overall assignment of confidence score for the genome annotation of *Geobacillus thermoglucosidasius* NCIMB 11955. The scores subsequently reflected, from the lowest scores, genes: not evaluated ("0"), included in the model but without literature or biochemical evidence ("1"), annotations stemming from genome annotation or from physiological data ("2"), evidence from genetic focused experimental approach ("3") or including extensive biochemical work on protein expression ("4").

In this project, the element of manual curation of each ORF has remained as the level of confidence in genome annotation was of great importance. The RAST SEED server was used to generate an annotation for comparison with that from ERGO Integrated Genomics [104] . Comparison of genome annotation to that of closely related species using PathwayBooster

has added another level of understanding of the metabolism which in turn became a backbone to generation of the metabolic model (as discussed in previous chapter). The steps involved in building the metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 can be roughly divided into initial reconstruction and model refinement. Whereas the former took a year, the latter has become an ongoing quest as experimental data was added.

4.2.1 Initial Reconstruction and Model Refinement

From genome annotation an initial metabolic network reconstruction can be relatively straight forward. Care needs to be taken when establishing the genome-derived metabolic capabilities of a system since some might be false and some simply omitted [134], which is why PathwayBooster assistance was a necessary step in weeding out the false positive metabolic reactions. It should be noted however that PathwayBooster operated mainly on KEGG pathways which in itself do not include all the reactions in the system, which in turn was one of many reasons for a scrupulous manual curation. PathwayBooster however was indispensable for immediate identification of missing reactions which were added to the metabolic reconstruction (however these reactions were added with a low confidence score). Whilst creating a genome-scale metabolic model one must be aware that the initial version of the metabolic model (even curated) will undergo constant improvements and additional experimental data will be incorporated into it, which means that a comprehensive data-mapping is essential. Over the period of this research, a constant improvement of genome annotation stemming from other studies has been incorporated into a metabolic model. Some of those findings were discussed already in previous sections and findings from the model were driving experimental focus on laboratory conditions.

When creating a database of metabolic functions, which is based on genome annotation, Enzyme Commission (EC) numbers were initially used to retrieve a metabolic reaction corresponding to a given gene. Databases such as BRENDA, EC2PDB, KEGG or MetaCyc were all used to get a corresponding reaction from EC numbers. The problem, however, with retrieving information about a given reaction from EC number, originates from their hierarchical nature (this is in detail described in Materials and Methods chapter). It was common to find in the genome of *Geobacillus thermoglucosidasius* NCIMB 11955 a gene with an EC annotation relating to enzyme class rather than a specific function (this has also been highlighted whilst assigning confidence score to annotations). This was however consolidated when analysing connectivity of a network and flux distribution using *in silico* gap-find

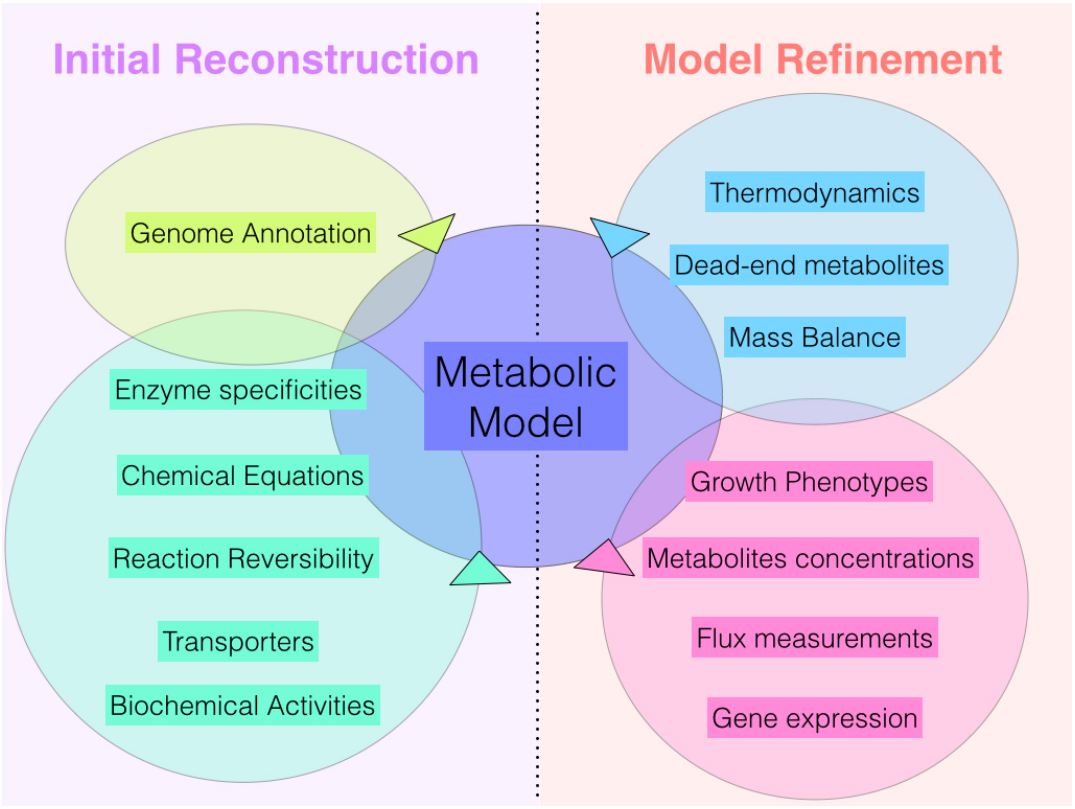


Figure 4.3: Model Refinement Stages.

methods from COBRA toolbox (discussed in detail in the next sections).

Building a reconstruction based on EC numbers may incorporate a few false positive reactions into a metabolic model [134, 49]. Assignments such as kinases, which are involved in signal transfer or DNA methyltransferases should not be involved in an *in silico* metabolic system [134] at this stage since they govern the kinetic models and are not involved in steady-state metabolic models. It should be noted that the refinement of the initial reconstruction was especially tricky given that most of the biochemical reactions, which originate from the genome annotations are not organism specific. A significant amount of time was allocated to combine literature confirmation and expert knowledge of this bacterium to confidently proceed with the metabolic model curation.

The reconstruction assembly, as recommended by numerous protocols available for the metabolic modeller were done pathway by pathway, using PathwayBooster, starting from central carbon metabolism and growing to the peripheral pathways where the correct annotation was less straightforward [134]. Reactions with lower confidence in both annotation and belonging to a certain pathway are annotated as such in a model. At this stage an essential gene-protein relationship (GPR) has been included in the reconstruction. Extra care was given when

describing genes encoding enzymes responsible for more than one reaction or if an enzyme has all the necessary subunits present. In the latter case, manual curation was applied and only reactions which were catalysed by an enzymes containing all the functional subunits were added to the reconstruction. At this stage information on subunits was added to the reaction when possible. Indeed, the gene-protein-reaction association was incorporated into the model with as much care and detail and possible. Enzymes often catalyse more than one reaction (one gene, multiple reactions), or multiple enzyme (and subsequently, more than one gene) are needed for carrying out one reaction (multiple genes, one reaction). In this case a Boolean "AND" and "OR" were introduced to the reconstruction to accommodate these instances.

The metabolic subsystems were categorised with agreement to the KEGG platform identification system. This was done in order to ease the process of finding reactions belonging to a given metabolic pathway or for parsing the information coded in the metabolic model.

The gaps in the initial reconstruction were verified using PathwayBooster with regards to known metabolic capabilities of evolutionarily closely related species (this has been discussed in a previous Chapter and in detail, the choice of strains is discussed in Materials and Methods chapter).

Generic reaction including metabolites with generic names have been excluded from this metabolic model or substituted for a more specific molecule. For example, these included reactions calling for metabolites as generic "alcohol", "protein" or "carbohydrate", where more than one specific metabolite belongs to a given classes. These reactions can be found in a majority of databases and often a decision based on literature had to be made about which of the reactions should be incorporated. The following "generic reaction" depicts the problem discussed (Reaction 4.1), where "A" represents a generic metabolite (such as "alcohol") and "A_1", "A_2" or "A_3" show potential candidates for a substitution. "P" stands for a generic product, such as "ester" whilst "P_1", "P_2" or "P_3" denote specific molecules. It is common that more than one of the substrates, or a cofactor can be utilised by a given reaction and such an instance is shown in reaction 4.4.

$$A + H_2O \rightarrow P \tag{4.1}$$

$$A_1 + H_2O \rightarrow P_1 \tag{4.2}$$

$$A_2 + H_2O \rightarrow P_2 \tag{4.3}$$

$$A_3 + B_1 \rightarrow P_3 \tag{4.4}$$

If a "generic reaction" is found such as (Reaction 4.1), all the instances of the reactions were considered (Reaction 4.1 - 4.4), especially when more than one version of a reaction is possible. Each instance of a reaction is then checked if the mass balance is conserved and considered as a potential candidate for a reaction in a model. This approach was based on a methodology suggested by [69, 70]. The above problem can be extended to cofactor usage by a given reaction. Special care was also applied when choosing between cofactors for given reactions in reconstruction of the metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955, especially where a reaction can be represented by a number of sub-reactions (as described above). The cofactor specificity in those instances where reconsolidated using an extensive literature search whilst was aided by used of PathwayBooster.

Reaction stoichiometry and a protonated state for each metabolite was calculated according to the protocol found in Chapter "Material and Methods". **Reaction directionality** was estimated for each reaction based on Gibbs free energy of group formation. This approach allows for estimation of reaction directionality based on $\Delta_f G'^o$ (free energy of formation) and $\Delta_r G'^o$ (free energy of reaction). This approach was complemented by a literature search and information stemming from network topology. This step is crucial for a falsely identified irreversible reaction may block a given metabolic flux and conversely, a false reversible reaction may give the system unrealistic metabolic capabilities and create too loose a constraint on a metabolic reconstruction [134].

The model of *Geobacillus thermoglucosidasius* NCIMB 11955 has been divided into two compartments; namely extracellular and intracellular with corresponding transport reactions. Transport reactions were created with accordance to metabolites present in the media and cell surrounding environment along with the need for secreted molecules.

Finally, **spontaneous reactions** were added at this stage, these included only reactions which could be connected to the set of metabolites already present in the reconstruction. These reactions were also assigned a mock gene and a protein to maintain gene-protein-reaction relationship hence making them easier identified during *in silico* analyses [134].

In steady-state metabolic models, reactions which are included in the network must be satisfying mass balance constraints and their net flux value must be zero. However, in the system there can be present molecules, which do not satisfy these requirements. Such molecules can be cofactors or lipopolysaccharides (LPS) [134]; although LPS can be quantitatively measured, the cofactors can't and hence are not included in the biomass reaction. In such

instances **demand reactions** are included in the reconstruction. In the model of *Geobacillus thermoglucosidasius* NCIMB 11955, there are only few such instances accounted for. Demand reactions, however are a very useful tool at the stage of model refinement, when they allow for a detection of gaps in the network by allowing substrate accumulation in the system. Demand reactions are set to be irreversible.

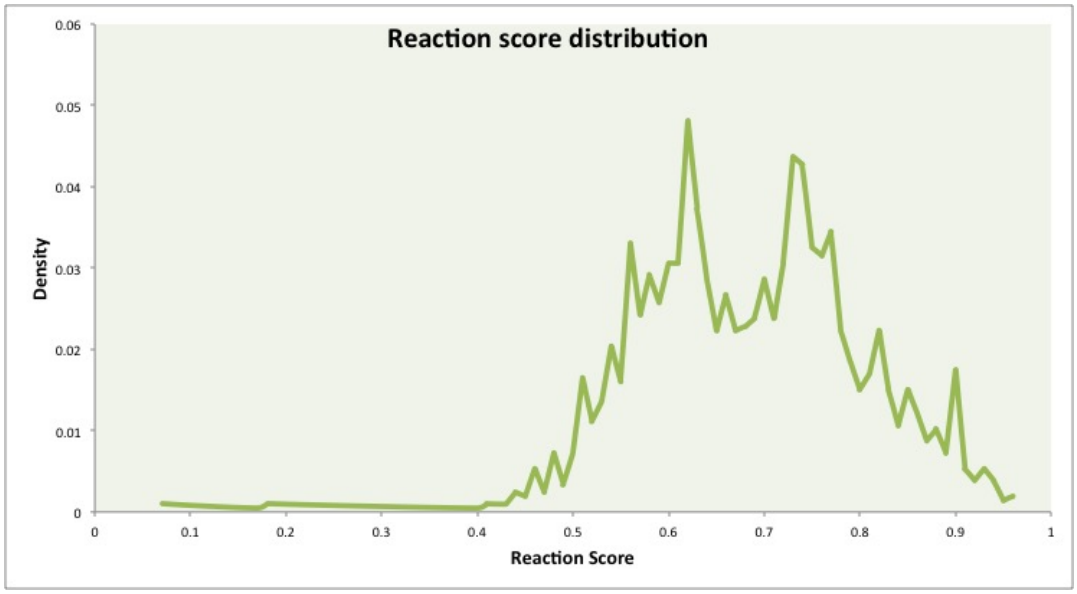


Figure 4.4: Check-met results: proportion of reactions in the model with score over 0.5 and below 0.5. CheckMets shows the extend of which the reactions are connected to the network and assigns higher score to reactions that are highly connected. The low-scoring reactions are those that are disconnected from the network 4.4.

Sink reactions on the other hand allow for reversible reaction when a metabolite may enter and exit the intracellular metabolite pool. These might include: molecules resulting from non-cellular processes that need to be accounted for in the network of reactions.

After the initial reconstruction and manual curation has been done, the network connectivity was checked using CheckMet tool [74].

4.3 Biomass, growth and non-growth associated maintenance requirements

4.3.1 Estimation of biosynthetic costs

A **growth-associated ATP maintenance reaction** was also added at this stage of reconstruction. This reaction describes ATP needed by the bacterium for metabolic maintenance, energy, which is necessary for the cell to replicate, allow for transport of molecules, synthesis of monomers or polymerisation and hence essential to account for in a metabolic model [134, 141]. Conversely, non-growth associated maintenance is an energy needed for maintaining functioning cell, excluding the energy required for growth [134, 141]. Initially, this reaction was estimated from literature and a "best guess" approach, which is based on macromolecule synthesis energy requirements [134]. In this methodology, the number of phosphate bonds is calculated per each macromolecule(DNA, RNA, protein) and accounted for in an ATP hydrolysis reaction, which is part of a biomass reaction (described in detail in the following section). The ATP hydrolysis reaction is incorporated as illustrated by Reaction 4.5, where "n" stands for the number of required phosphate bonds as calculated from experimental and theoretical methods.

$$nATP + nH_2O \rightarrow nADP + nP_{\mathrm{i}} + nH^+ \tag{4.5}$$

This approximation which always results in a much lower value than an actual system requirement should display has also been reconsolidated with experimental data. The actual growth-associated ATP maintenance was calculated from chemostat growth experiments (see Figure 4.5 [121] and Materials and Methods chapter where theoretical approach is detailed). The GAM was estimated from a plot estimating glucose uptake rate against growth rate (Figure 4.5b). The GAM was estimated to be **33** (which aligns with the experimental data).

The relationship between the growth rate, substrate uptake and ATP consumption is expressed via the Eq 4.6 which was used to calculate the appropriate values from experimental data. In Eq 4.6: Y_{atp} stands for ATP yield on glucose (mmol · mmol $^{-1}$), C_{sub}: rate of glucose uptake(mmol · gDCW · h $^{-1}$), Y $_{max}$: growth on glucose (excluding GAM and NGAM) [58, 134].

$$Y = \frac{Y_{\text{atp}}}{GAM + \frac{Y_{\text{atp}}}{Y_{\text{max}}}} - \frac{NGAM}{(GAM + \frac{Y_{\text{atp}}}{Y_{\text{max}}}) \cdot C_{\text{sup}}} \tag{4.6}$$

The growth-associated maintenance requirement was estimated from weighted least square regression to be **25.44** ATP mmol/gDCW/h. The y-intercept (0.1551 mol glucose/gDCW/h) allows for the estimation for the **non-growth associated maintenance**, which in this case can be calculated to **1.53** mmol ATP/gDCW/hr.

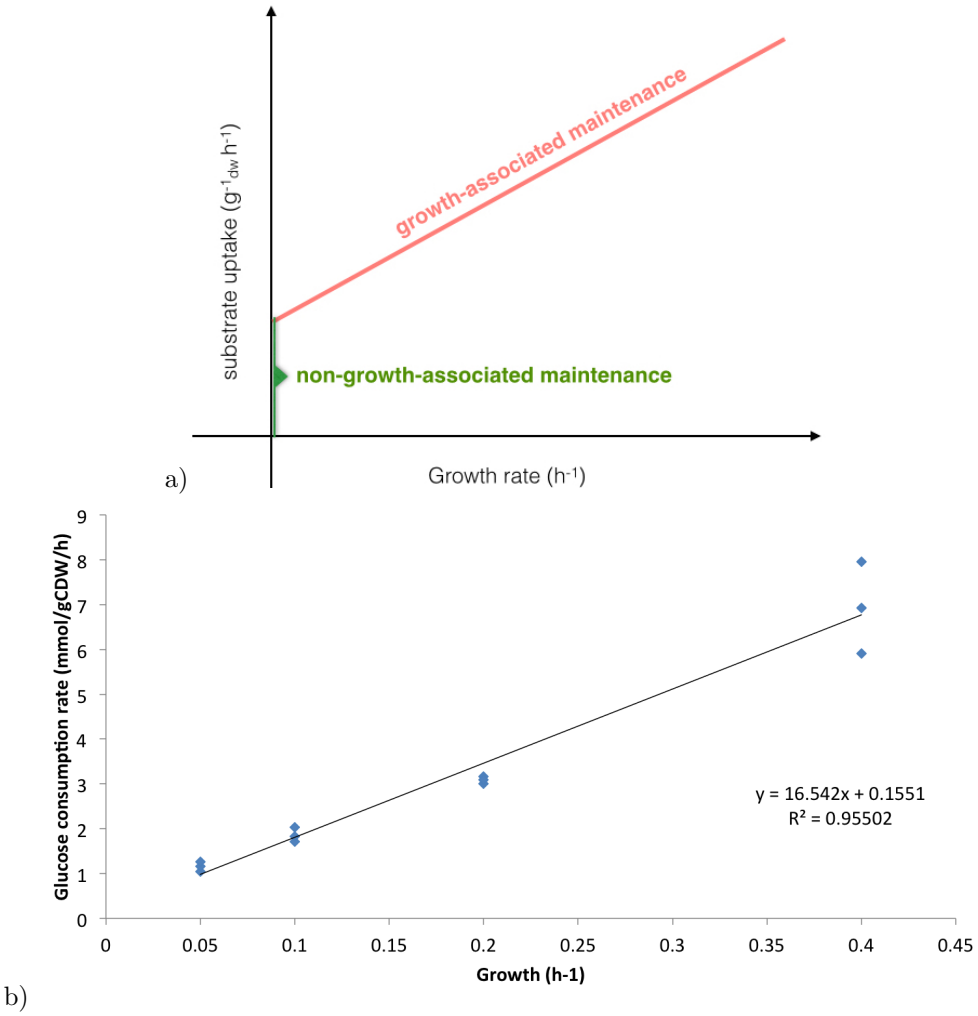


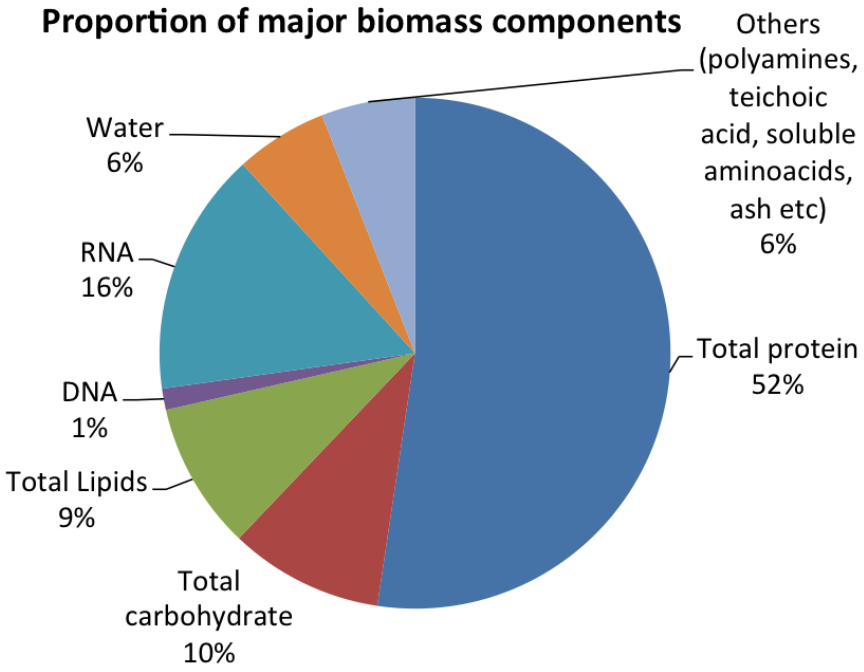
Figure 4.5: Estimation of growth-associated maintenance(GAM) and non-GAM from chemostat growth experiments: a) theoretical, b) experimental estimation.

4.3.2 Growth requirements

The growth requirements, such as rich and minimum growth media composition for *Geobacillus thermoglucosidasius* NCIMB 11955 were estimated from prior experiments (and can be found in the Materials and Methods chapter). This bacterium has been a focus of scientific research for over a decade and so the growth requirements are in turn well defined.

4.3.3 Biomass

A biomass objective function must be defined for a metabolic model to correctly describe cell growth requirements. It is needed to correctly simulate the metabolic capabilities of a given microbial system and allows for model simulation using a range of Constraint-Based Reconstruction and Analysis (COBRA) methods [50]. To answer a variety of questions relating to a microbial system, it is important to make the biomass objective function as physiologically realistic as possible [50]. Correctly calculated biomass is crucial in flux balance analysis and encompass rates and proportions of its components [50] .



	% proportion	% of dry mass	umol/g CDW			
			R1A1	R1A2	R1An1	R1An2
Asp/Asn	9.822901778	5.143271371	470.9439185	456.9380975	543.2945417	445.2021817
Thr	5.223629803	2.735092565	285.1531998	276.6727746	328.9609885	269.5667609
Ser	3.032535947	1.587835822	192.1825435	186.4670555	221.7073473	181.6778693
Glu/Gln	15.9866549	8.370612505	683.2743563	662.9538512	788.2450833	645.9266639
Gly	4.471915145	2.34149477	432.5595756	419.6953008	499.0132522	408.9159223
Ala	7.02518988	3.678389421	545.5404408	529.3161273	629.3512499	515.7212673
Val	7.165843244	3.752035523	399.0794279	387.210849	460.3895843	377.2657954
Met	2.142098805	1.121602935	90.15616617	87.47493157	104.006764	85.2282412
Ile	5.931861476	3.105922669	289.4274972	280.8199548	333.8919418	273.6074257
Leu	7.555727299	3.956178814	368.6591891	357.6953048	425.295916	348.5083231
Tyr	3.249366635	1.70136837	109.9102876	106.6415676	126.7956905	103.9026048
Phe	4.104885016	2.149317794	153.9610604	149.3822752	177.6139376	145.5455677
His	2.051700609	1.074270439	82.56968965	80.11407604	95.25478498	78.05644055
Lys	7.811917709	4.090320113	336.4921384	326.4848987	388.1870748	318.0995195
Arg	5.496038973	2.877726006	194.2461579	188.4692982	224.0879926	183.6286867
Pro	3.10850815	1.627614867	176.6881844	171.433497	203.832606	167.0304299
Trp	3.677125826	1.925343082	108.8819892	105.6438506	125.6094156	102.930513
Cys	2.142098805	1.121602935	114.4424202	111.0389151	132.0240899	108.1870116

DNA			mmol/ gDCW	mmol/ gDCW	mmol/ gDCW	mmol/ gDCW
Nucleotide			R1A1	R1A2	R1An1	R1An2
dAMP			0.029903043	0.016155694	0.024957341	0.020698191
dCMP			0.023333636	0.012606446	0.019474457	0.016151001
dTMP			0.029903043	0.016155694	0.024957341	0.020698191
dGMP			0.023333636	0.012606446	0.019474457	0.016151001
Total			0.106473358	0.057524279	0.088863598	0.073698385

	Proportion	MW g/mol	R1A1	R1A2	R1An1	R1An2
Arabinose	6.849713082	150.13	0.889335902	0.780563103	0.509537249	0.939296505
Galactose	38.23636232	180.16	4.964437102	4.357247268	2.844330941	5.243326407
Glucose	1.932156606	180.16	0.250862513	0.220180048	0.143729489	0.264955324
Xylose	30.68206581	150.13	3.983621261	3.496392945	2.282381059	4.207410855
Mannose	1.154935716	180.16	0.149951653	0.131611382	0.085913492	0.158375551
Fructose	21.14476647	180.16	2.745341262	2.409564361	1.572919333	2.899567472

			% of dry mass		
Component	Proportion	R1A1	R1A2	R1An1	R1An2
Glycerol	12.32876712	1.388977266	1.214529841	1.329814944	1.412515947
Fatty acid	87.67123288	9.877171673	8.636656651	9.456461822	10.04455784

Figure 4.6: Breakdown of biomass composition for *Geobacillus thermoglucosidasius* NCIMB 11955. The experimental estimations of biomass components were analysed by Dr Shyam Masakapalli. The table present the raw estimates of biomass components from 8 replicates. The suffiix "An" denotes anaerobic cultures.

The objective function and growth requirements described in the previous section allow for a feasible simulation of bacterial behaviour. It is most likely that a given microorganism does not live in an environment rich in nutrients but rather grows in a nutrient-limited habitat [50]. Data needed for the description of habitat richly supplemented or on minimal nutrient can be estimated from batch and chemostat cultures. The biomass composition for the metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 was initially estimated theoretically and later the values were corrected to those gathered by the experimental data. In an overview, the biomass components can be divided into 9 areas(as shown by the figure 4.6). The following tables show calculated experimentally or estimated biomass composition for *Geobacillus thermoglcusoidasius* NCIMB 11955. Tables: 4.2, 4.3,4.4 , and 4.5 were estimated from experiments and respectively show: amino acids, DNA, RNA, lipid and fatty acids. The values shown in tables: 4.1 and 4.6 were estimated according to the protocol found in chapter 8.2.7. The experimentally validated amounts of RNA, DNA, lipid and fatty acids were measured by Dr Shyam Masakapaclli.

Name	Mol. Weight (g/mol)	Proportion(%)	µmol/g of CDW
K	39.1	88.4	708.2
Mg	24.3	6.2	98. 8
Fe^{+3}	55.9	0.7	4.1
Ca	40.1	0.35	3.15
Phosphate	96.0	3.7	17.3
Diphosphate	174.9	0.7	1.2
Sum		100.05	

Table 4.1: Ion composition for biomass of metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 from paper by Tang *et al.*, 2009 [130]

Name	Mol. Weight (g/mol)	Proportion(%)	Average (mg/g)	µmol/g of CDW
Glycine	57	4.47	20.73	456.63
L-Alanine	71	7.03	32.56	575.89
L-Valine	99	7.17	33.21	421.28
L-Leucine	113	7.56	35.02	389.17
L-Isoleucine	113	5.93	27.49	305.53
L-Serine	87	3.03	14.06	202.87
L-Threonine	101	5.22	24.21	301.02
L-Phenylalanine	147	4.10	19.03	162.53
L-Tyrosine	163	3.25	15.06	116.03
L-Tryptophan	186.2	3.68	17.04	114.94
L-Cysteine	103.2	2.14	9.93	120.81
L-Methionine	131	2.14	9.93	93.17
L-Lysine	128	7.81	36.21	355.21
L-Arginine	156	5.50	25.47	205.05
L-Histidine	137	2.05	9.51	87.16
L-Aspartate	115	4.91	45.53	248.575
L-Glutamate	129	7.995	74.10	360.645
L-Asparagine	115	4.91	45.53	248.575
L-Glutamine	129	7.995	74.10	360.645
L-Proline	97	3.11	14.41	186.52
Sum		100	463.50	

Table 4.2: Amino acid composition for biomass of metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 as established through experimental approach.

Name	Mol. Weight (g/mol)	Proportion(%)	g/mol	mmol/g of CDW
DNA				
dAMP	313.2	28.1	87.962	0.299
dCMP	289.2	21.9	63.378	0.233
dTMP	304.2	28.1	85.435	0.299
dGMP	329.2	21.9	72.144	0.233
Sum		100	308.919	1.06
RNA				
AMP	113	26.6	87.567	0.1132
GMP	87	34.3	118.404	0.1460
CMP	101	18,8	57.378	8.00
UMP	147	20.4	62.465	8.68
Sum		100	325.813	16.9392

Table 4.3: DNA and RNA composition for biomass of metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 as established through experimental approach.

Name	Mol. Weight (g/mol)	Proportion(%)	Average (mg/g)
Phosphatidylglycerol	722.969862	25	0.1561
Diphosphatidylglycerol	1466.0585	50	0.3862
Phosphatidylethanolamine	271.161722	25	0.1365
Sum		100	

Table 4.4: Main lipid composition for biomass of metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 as estimated from the contribution from fatty acids and phospholipids

Name	Mol. Weight (g/mol)	Proportion(%)
iso-C14:0	228.37	0.4
n-C14:0	228.47	2.5
iso-C15:0	242.4	9.6
Anteiso-C15:0	242.4	3.3
iso-C16:0	256.42	27.9
n-C16:0	256.42	34.7
iso-C17:0	270.45	11.4
Anteiso-C17:0	270.45	6.1
n-C18:0	284.48	4.3
Sum		100.2

Table 4.5: Fatty acid composition for biomass of metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 from paper by Tang *et al.*, 2009 [130]

Name	Mol. Weight (g/mol)	Proportion(%)	µmol/g of CDW
Menaquinol 7	651.0	1.0	0.3
10-Formyltetrahydrofolate	471.4	1.0	0.4
NAD	662.4	61.9	16.2
NADP	345.2	9.3	4.7
NADPH	503.2	8.7	3.0
AMP	424.2	6.3	2.6
ADP	321.2	1.9	1.0
ATP	740.4	4.0	0.9
CMP	479.1	1.5	0.5
GMP	361.2	1.1	0.5
CDP	519.1	1.3	0.4
GDP	400.2	0.6	0.3
CTP	741.4	0.9	0.2
GTP	440.2	0.5	0.2
Sum		100.2	

Table 4.6: Metabolite content for biomass of metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 from paper by Tang *et al.*, 2009 [130]

4.4 Flux Balance Analysis

The refined model includes 1,011 reaction associated with 859 genes, out of which 950 have calculated Gibbs free energy change (as discussed in previous sections) and 419 reactions were found to be irreversible according to this methodology (41.4%). The model was validated and blocked reactions found according to COBRA toolbox protocol [61], which are also described in detail in Materials and Methods chapter. The gaps in the model were analysed with COBRA toolbox methodology using such tools as GapFind() [61] to assist in finding gaps in the metabolic network or detectDeadEnds()[61] t that helps detect dead-end metabolites. The FBA was used also in the preliminary stages for model refinement purposes to check if the model can predict fluxes in the central carbon metabolism and produce core metabolites using biomass objective function. FVA analysis on rich media have returned 128 reactions where both the minimal flux and the maximum flux were "0"(12.6%). These were reactions found on the peripheral pathways of the metabolic network.

Fluxes given by the model were compared to those presented by Tang *et al*, 2008 paper. It should be however noted that the flux distribution described by Tang *et al*,2009 show central carbon metabolism for *Geobacillus thermoglucosidasius* M10EXG and only refer to fluxes derived from ^{13}C flux model and encompassed only: citric acid cycle, glycolysis, pentose phosphate pathway, ED pathway (absent in *Geobacillus thermoglucosidasius* NCIMB 11955) and thus does not account for an entire spectrum of reactions provided by genome-scale metabolic model. Tang *et al*, 2008 model does not include glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.59) which also contributes to the fluxes in glycolysis. ^{13}C flux model and genome-scale metabolic models, although both in a steady-state, each gives a different depth and insight into bacterial metabolism. The glyceraldehyde-3-phosphate dehydrogenase (NADP dependents) shows that a third of a flux under aerobic conditions runs through the NADP-dependent GAPDH, whilst the remaining two thirds are catalysed by NAD-dependent GAPDH.

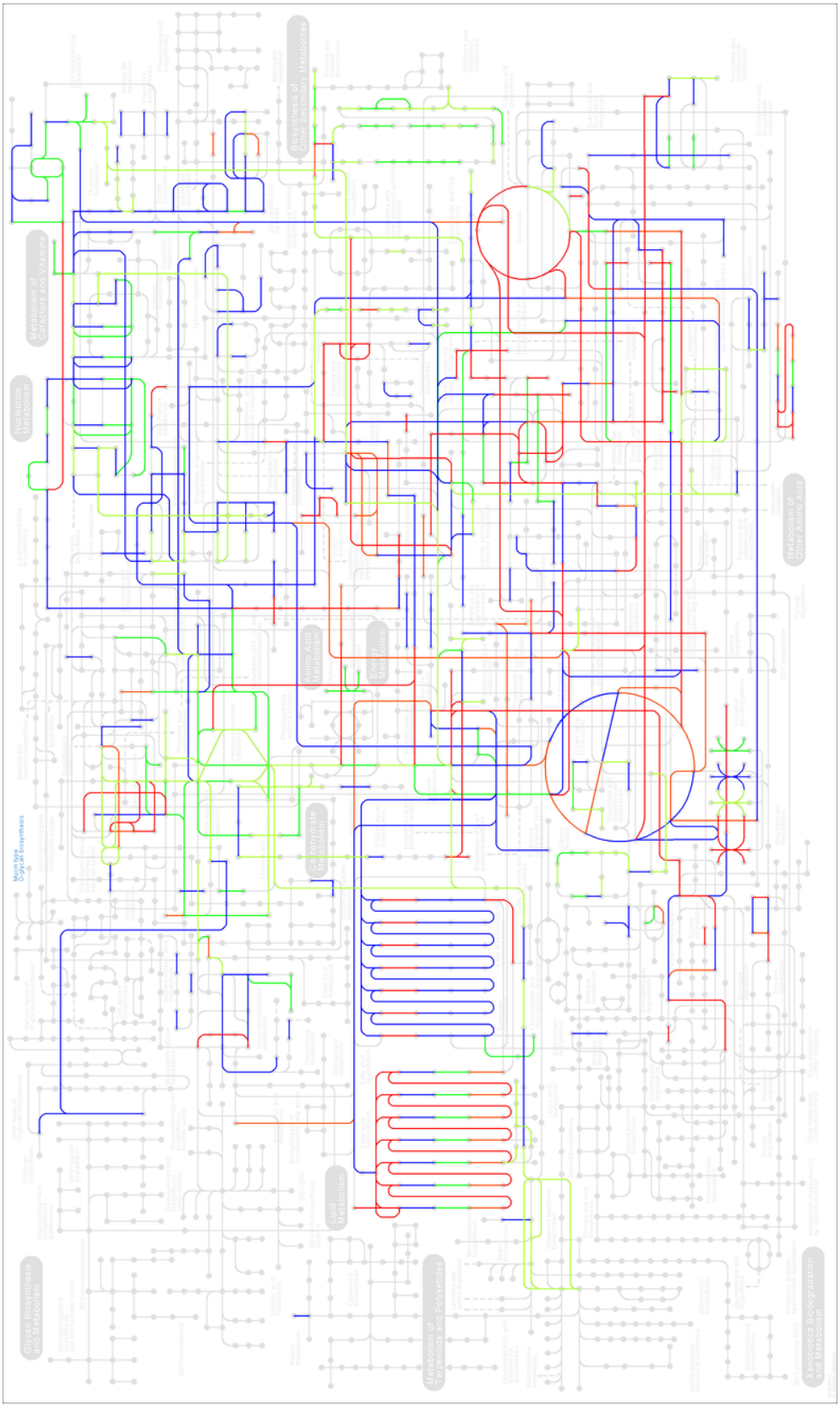


Figure 4.7: Glucose flux distribution for minimal against rich media growth across the genome-scale metabolic reconstruction of *Geobacillus thermoglucosidasius* C56-YS93. The following colour denote: blue: no change in flux, GreenYellow: significant change in flux but not 2^2 fold increase, Green: over 2^2 fold increase, OrangeRed: significant but less than 2^2 fold decrease and Red: over 2^2 fold decrease.

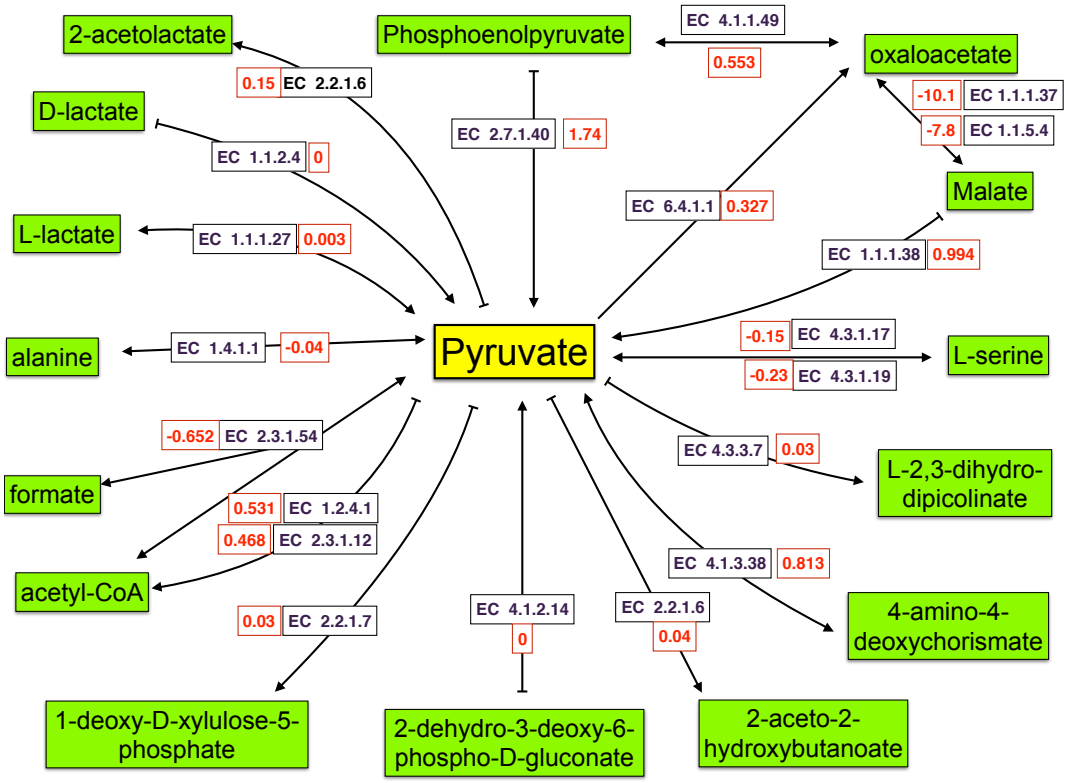


Figure 4.8: Pyruvate fluxes distribution for metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955. The negative value suggests consumption for reversible reactions. The reactions corresponding to enzyme numbers (EC..) can be found in the supplementary data, along with reactions they represent.

4.5 Conclusions

"The whole is greater than the sum of its parts" Aristotle.

In this chapter, the reconstruction of the *Geobacillus thermoglucosidasius* NCIMB 11955 model has been explained in detail. The reconstruction benefits from a vast amount of manual curation from early stages of reconstruction and incorporates experimental data in estimation of bacterial requirements, such as biomass or environmental conditions. This reconstruction is the first manually created genome-scale metabolic model for any species within the genus *Geobacillus*. This model is an up-to-date representation of genomic and biochemical knowledge available for this strain. The FVA findings suggest that, under aerobic conditions, there is no minimum or maximum flux going through 2-oxoglutarate synthase. We have also found that GAPDH, described in the previous section carries a flux from glyceraldehyde-2-phosphate in both aerobic and anaerobic environments. The data on maximum and minimum fluxes (FVA analysis) through the network of reactions is provided in the supplementary data.

Although metabolic models have great potential in understanding capabilities of a system, their simplification shows a significant limitation on their predictive capabilities [47, 85]. The view of a cellular system from the fluxes perspective does not include the intricate regulation that goes on in the cell. From transcriptional , gene or protein regulation to precise enzyme kinetics, a vast amount of "omics" data is needed to refine the model and make it a true representation of whole-cell metabolism [92]. Transcriptional regulation in a given condition for example can affects fluxes by down-regulation of certain pathways and directing the major flow in a different direction. This in turn leads to a false metabolic representation of the fluxes in steady-state modelling [47].

Whilst genome-scale metabolic models are providing a sufficient prediction of fluxes, incorporating the models with gene expression data provides an added bonus to its predictive powers [92]. However, the other possibility to overcome the problem of capturing the nature of different metabolic states under specific transcriptional regulation, is to build multiple models with each describing a desired state on its own. The resulting models are still based on a steady-state constraint-based modelling approach but shed light onto a difference in flux distribution when compared to the "main" model.

The FBA and FVA approach is a very useful tool in validating the quality of the predicted metabolic functions described in the previous chapters. The predicted annotation of the

2-oxoglutarate synthase for example is predicted to be inactive in the the bacterial *in silico* growth in both aerobic and aerobic conditions, on rich and minimal media alike. This might be an indication that this is indeed an inactive artefact in this strain however definite conclusions cannot be made without experimental validations. The FBA is a useful tool in analysing the fate of a given metabolite. For example, in the Figure 4.8 the complete collection of fluxes is shown for pyruvate. This information is paramount when combined with the observation of bacterium in laboratory conditions. Such information can answer the questions such as what might be the reasons for impaired growth under anaerobic conditions (explained in the next chapter). The simulations of metabolic models can in fact provide the researcher with a detailed insight into the metabolic network and as such help uncover new and creative strategies and designs in strain engineering.

Chapter 5

Predictive modelling and genome-scale metabolic model analyses

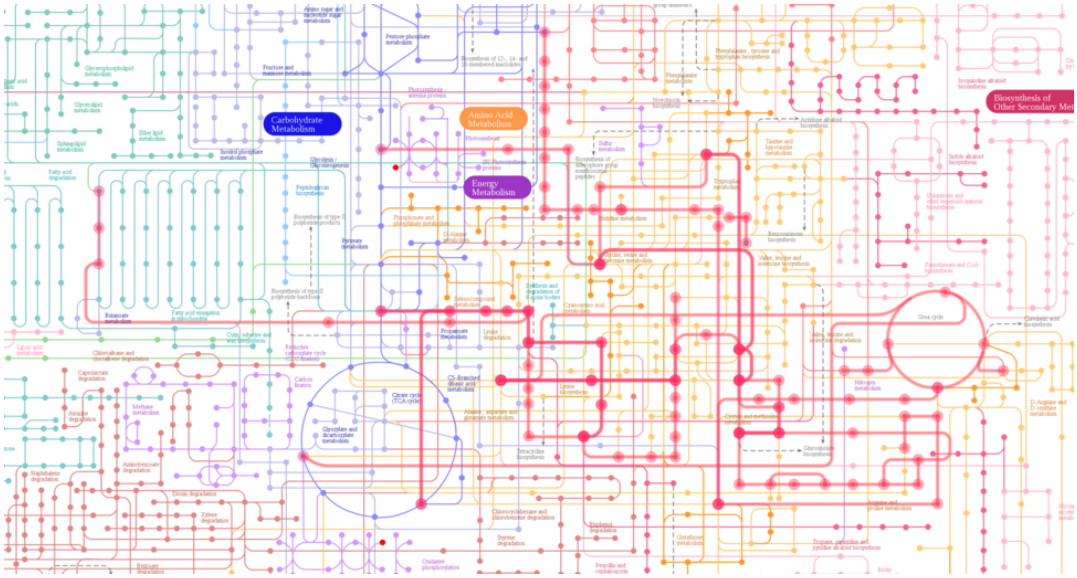


Figure 5.1: Representation of global metabolism of *G. thermoglucosidasius* NCIMB 11955

5.1 Introduction

The metabolic model is a powerful tool, which can provide a novel an insight into the metabolic capabilities. The properly defined constraints and conditions should allow for the metabolic behaviour to be modelled under different environmental perturbations. In this chapter, the model flux landscape is discussed, with special focus on perturbed metabolism in knock-out strains. The flux distribution landscape will be first discussed in light of flux balance analysis (FBA), as briefed in the previous chapter. The FBA metabolic profiles will be shown for aerobic conditions and under anaerobic conditions. For this Minimisation of Metabolic Adjustment (MOMA) will be performed as it has been shown to more accurately predict perturbed metabolic network's flux distribution rather than FBA (see Figure **??**) [115]. In essence, flux balance analysis assumes optimal metabolic adaptation for growth under various conditions by changing the topology of the metabolic network, which works well for normal predictions. MOMA, on the other hand allows for prediction of perturbed metabolic networks such as knockout growth rates. This results in a much more accurate flux distribution than that of FBA.

This chapter will finish with *in silico* gene knockouts for production of succinate, along with corresponding flux redistribution and gene knock-ins for butan-2,3-diol production.

5.2 Succinate Production

Succinate is an important compound which can be used in agriculture, food and pharmaceutical industries. Succinate has been identified as a promising substrate for production of polybutyrate succinate (PBS) or polyamides (such as Nylon) or green solvents [122, 26]. The production of succinic acid by bacterial fermentation has been a focus for over a decade, which provides an alternative to costly production of this compound from liquified petroleum gas or oil [122]. *Geobacillus thermoglucosidasius* NCIMB 11955 as a proven fermentative microorganism could be a potentially useful factory for succinate production due to it's ability to grow under both aerobic and anaerobic conditions and to withstand high temperatures. Succinate is a natural product of anaerobic growth and is used as an intermediate in the tricaboxylic acid cycle (TCA). *G. thermoglucosidasius* in chemostat conditions yields 3.618 mmol of succinate, 3.11 mmol of ethanol and 5.092 mmol of acetate per 1 g per gram dry cell weight (DCW) when supplied with 0.125 mmol of glucose. *Geobacillus thermoglucosidasius* can produce succinate through the following routes:

- R00405: succinate:CoA ligase (ADP-forming) (Citric acid cycle)
- R02164: succinate:quinone oxidoreductase (Citric acid cycle)
- R00479: isocitrate glyoxylate-lyase (succinate-forming) (Glyoxylate and dicarboxylate metabolism)
- R00713: succinate-semialdehyde:NAD_+ oxidoreductase (Alanine, aspartate and glutamate metabolism)
- R00714: succinate-semialdehyde:$NADP_+$ oxidoreductase (Alanine, aspartate and glutamate metabolism)
- R00999: O-Succinyl-L-homoserine succinate-lyase (Cysteine and methionine metabolism)
- R01288: O-succinyl-L-homoserine succinate-lyase (Cysteine and methionine metabolism)
- R02508: cystathionine gamma-synthase (Cysteine and methionine metabolism)
- R03260: O-Succinyl-L-homoserine succinate-lyase (adding cysteine) (Cysteine and methionine metabolism)
- R10343: succinyl-CoA:acetate CoA-transferase (Citric acid cycle)

Two alternative ways of *in silico* succinate production were considered: either through knock-out of genes in order to direct flux toward succinate or through inserting new genes into the system. Using OptKnock analysis [42] the options for deleting a number of genes to increase the flux towards succinate production, knock-outs of one to four genes was considered. OptKnock was used to find the potential candidates for the *in silico* strain engineering. MOMA was used to simulate the fluxes of the perturbed network.

5.2.1 Knock-ins approach for production of the succinate

One of the incomplete pathways present in *Geobacillus thermoglucosidasius* NCIMB 11955 is a route leading from L-glutamate to succinate as shown in Figure 5.2. In this route the missing enzyme is glutamate decarboxylase (EC 4.1.1.14). The glutamate decarboxylase catalyses the conversion of L-glutamate to 4-aminobutanoate. Glutamate decarboxylase is found in *Bacillus megaterium*, which makes this knock-in a viable candidate for further experimental testing in laboratory conditions. The flux distribution was evaluated after addition of the reaction 5.1. The result of model simulations can be viewed in Table 5.1. The knock-in of a gene and hence reaction associated with it into the network increases the theoretical succinate production by 36%, although a slight change in growth can be observed. It should be noted that this is a succinate producing hypothetical strain, only under micro-aerobic conditions as succinate is already naturally produced under those conditions.

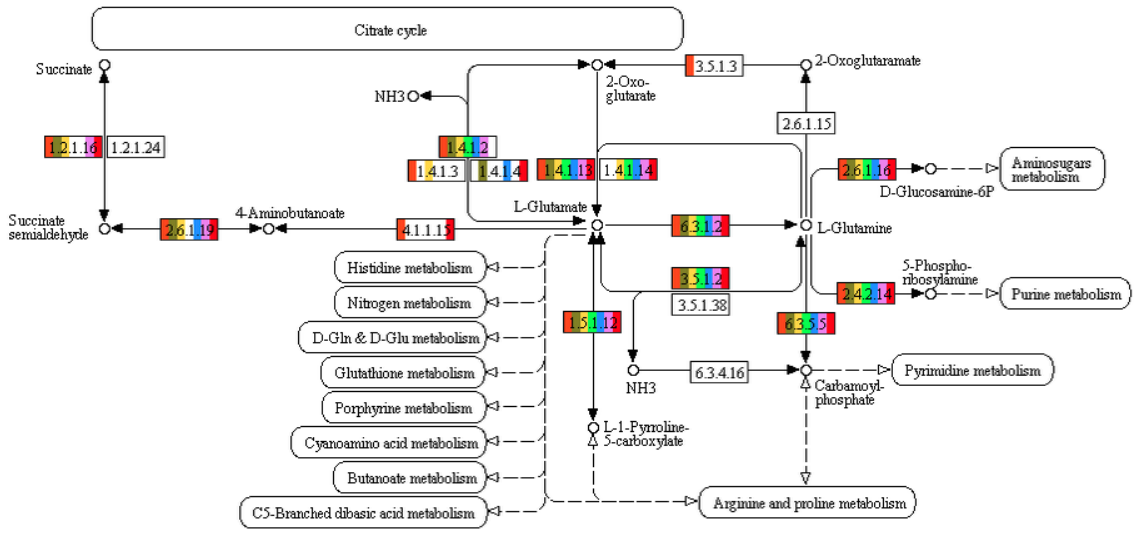


Figure 5.2: Incomplete route of succinate production in *Geobacillus thermoglucosidasius* NCIMB 11955 from L-glutamate.

$$L - glutamate \rightarrow 4 - aminobutanoate + CO_2 \quad (5.1)$$

Model	Biomass(h^{-1})	% Biomass	Succinate (mmol/g of DW h^{-1})
WT	0.42420912	100	0.21371425
Glu_KI	0.37266571	87.84	0.2926412
	% Succinate	Ethanol (mmol/g of DW h^{-1})	Acetate (mmol/g of DW h^{-1})
WT	100	15.6	13.863
Glu_KI	136.93	16.7	13.164

Table 5.1: Simulations of the wild type strain of *Geobacillus thermoglucosidasius* NCIMB 11955 against a mutant strain with knock-in of glutamate decarboxylase. Glu_KI stands for a knock-in strain and WT for wild type strain. Glucose uptake rate was 16 mmol/ g DCW h^{-1}.

5.2.2 Knock-out approach to succinate production

The other *in silico* approach to increase in the succinate production was to find candidate reactions in a knockout approach. MOMA was used in as before in evaluating fluxes in the perturbed network. OptKnock was used to identify candidate knock-out reactions. The key knock-out methodologies relied on ultimate overexpression of glyoxylate cycle (Figure 5.3. The glyoxylate cycle is a series of anaplerotic reactions, that differs from the citric acid cycle in that it bypasses the reactions, which lead to the the loss of CO_2. The knock-out approach identified through *in silico* model simulation has focused on the improved production of acetyl-CoA, which joins the glyoxylate cycle and ultimately improves the theoretical yield of succinate production. The following genes were knocked-out in order to achieve this aim; namely pyruvate kinase and malate dehydrogenase. Table 5.2 show the results of model simulation. As the theoretical results suggest, such double knock-out can increase succinate production significantly, however, it also dwarfed the growth by more than a half. This knock-in strategy for succinate production is overall an route worth pursuing in the experimental even though the growth of the bacterium was decreased to 41% of the theoretical wild-type mutant. This *in silico* can be overcome by engineering inducible promoter to yield higher succinate flux even at half the growth of the strain.

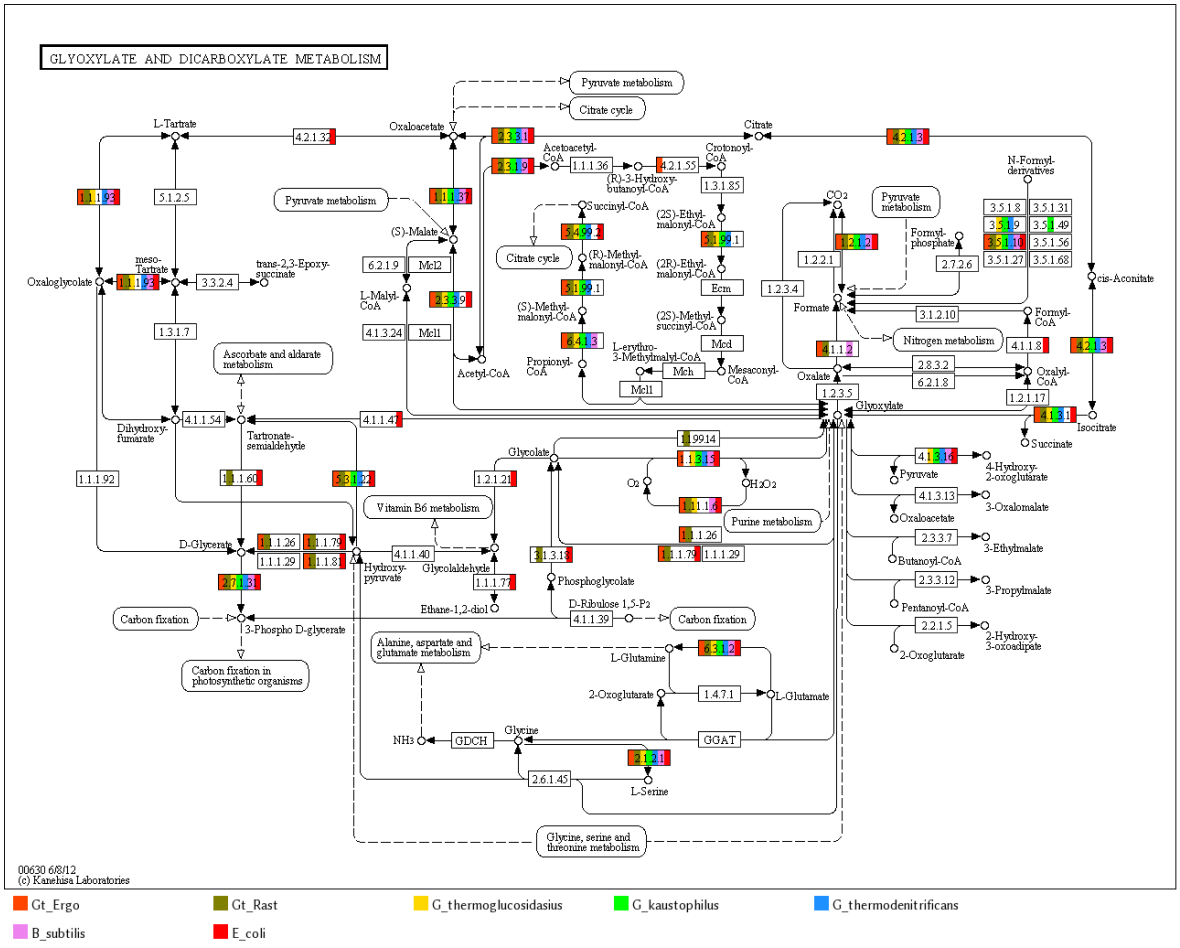


Figure 5.3: Glyoxylate cycle as annotated for the species in genus *Geobacillus*.

Model	Biomass(h^{-1})	% Biomass	Succinate (mmol/g of DW h^{-1})
WT	0.42420912	100	0.1068
PykMdh_KO	0.17682192	41.46	0.483

	% Succinate	Ethanol (mmol/g of DW h^{-1})	Acetate (mmol/g of DW h^{-1})
WT	100	15.38	14.863
PykMdh_KO	452.03	16.271	14.288

Table 5.2: Simulations of the wild type strain of *Geobacillus thermoglucosidasius* NCIMB 11955 against a mutant strain with pyruvate kinase and malate dehydrogenase knock-out .PykMdh_KO stands for a knock-out strain and WT for wild type strain. Glucose uptake rate was 16 mmol/ g DCW h^{-1}

5.3 Butan-2,3-diol Production

Butan-2,3-diol is a biochemical with an important role in the biofuel industry. What makes butene-2,3-diol such an important candidate is its heating value that is comparable to that of other liquid fuel (27,200 J/g) [29, 62]. Butane-2,3-diol can also serve as a substrate for production of chemicals such as methylethylketone, gamma-butyrolactone and 1,3-butadiene [29]. Microbiological production of this chemical has been observed in bacterial strains such as *Klebsiella pneumoniae* or *Paenibacillus polymyxa* [62].

The requirement for countries to secure renewable sources of biofuels with low carbon emission has gathered momentum over the past years and, is especially relevant given the agreements of the United Nations Climate Change Conference (Paris, 2015) [137] and Sustainable Developmental Goals [138]. It is now necessary for countries to seek to change their economies to be pro-environment.

5.3.1 Butan-2,3-diol *in silico* production

One of the routes for production of butan-2,3-diol is from pyruvate. This route is shown in Figure 5.4. The steps are carried out by the acetolactate synthase (EC 2.2.1.6), acetolactate decarboxylase (EC 4.1.1.5) and D-butanediol dehydrogenase (EC 1.1.1.4). These enzymes catalyse the pyruvate as a substrate, that is then converted to butan-2,3-diol through 2-acetolactate and 2-acetoin. In the majority of the species in the genus *Geobacillus*, from the gene assignment point of view, this pathway seem incomplete, lacking D-butanediol dehydrogenase and acetolactate decarboxylase (see Figure 5.4). However, the gene encoding D-butanediol dehydrogenase has been found within the genome of *Geobacillus themoglucosidasius* NCIMB 11955, which make this species a potential producer of butan-2,3-diol if the route was completed with acetolactate decarboxylase.

Using the same methodology as in the previous section, the flux distribution was modelled for a mutant strain with acetolactate decarboxylase a gene knock-in. Acetolactate decarboxylase catalyses the reaction 5.2.

$$(S) - 2 - Acetolactate \leftrightarrow (R) - Acetoin + CO_2 \tag{5.2}$$

The results predicted by the *in silico* metabolic model simulations suggest that the growth is impacted only by a marginal level, whilst the production of butan-2,3-diol is on a similar

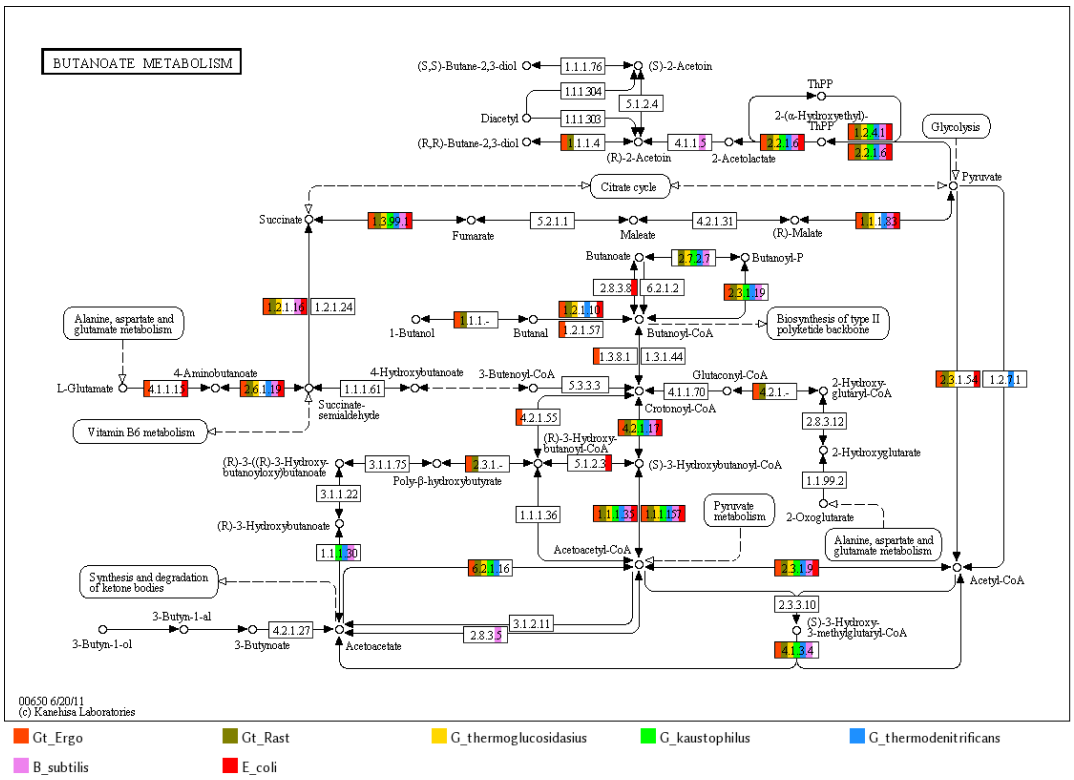


Figure 5.4: Butanoate metabolism pathway with the gene assignments for the species in genus *Geobacillus*.

level to that for succinate production, after double mutation, as discussed in the previous section.

Model	Biomass(h^{-1})	% Biomass	Butan-2,3-diol (mmol/g of CDW h^{-1})
WT	0.42420912	100	0
But_KI	0.42412409	99.97	0.4356

Table 5.3: Simulations of the wild type strain of *Geobacillus thermoglucosidasius* NCIMB 11955 against a mutant strain with acetolactate decarboxylase knock-in. But_KI stands for a knock-in strain and WT for wild type strain. Glucose uptake rate was 16 mmol/ g DCW h^{-1}.

5.4 Discussion

The *in silico* predictions using the genome-scale metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955 demonstrate the potential applications of this model for strain improvement. This microorganism is an important platform that is already used for the production of ethanol. Based on the *in silico* analyses presented here, this bacterium can be used also as a platform for the production of succinate or butan-2,3-diol. However, the model is a representation of metabolic pathways, as deduced from the gene assignment and does not account gene regulation and hence, the behaviour of those metabolic reactions [35]. It has been shown that the problem with simulating correct flux distribution for knock-out strains is often limited due to the gene regulation not being incorporated [35]. The model could be further curated to display a more accurate representation of the metabolism if gene expression data is used to constrain the fluxes. This can be achieved using mathematical methodologies, such as E-Flux [35] , iMAT [151] ,GIMME [13] , DPA [17] or PRIAM [32] (the rationale of these approaches is presented on an example of E-Flux in Figure 5.5). The gene expression data for each gene is used to constrain the maximum flux for the corresponding reaction and ultimately the qualitative power of the model can be increased.

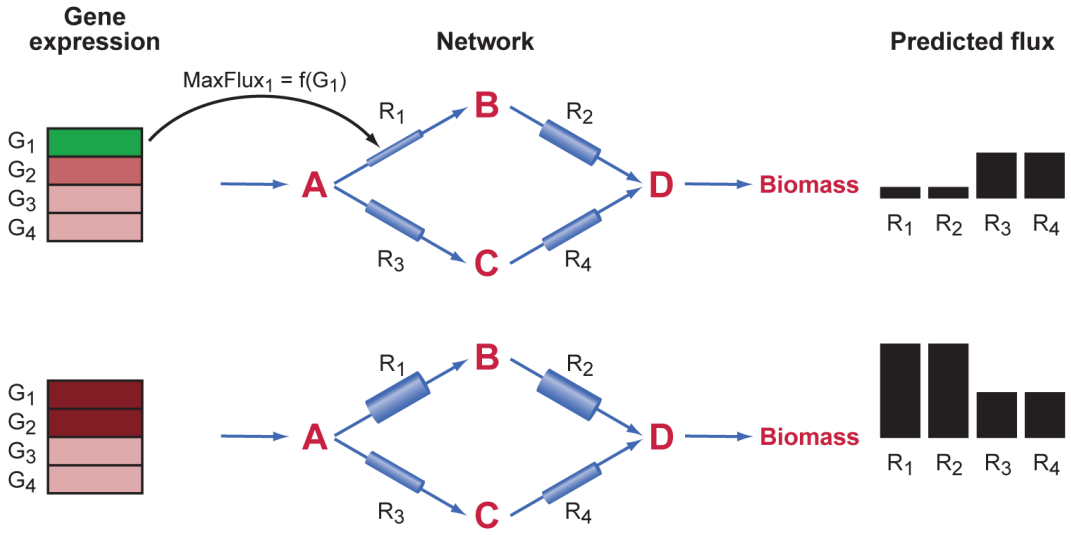


Figure 5.5: The conceptual overview of E-Flux approach to incorporating gene expression data to genome-scale metabolic modelling. Diagram taken from the original paper by Colijn *et al*, 2009 [35].

Chapter 6

Oxygen-limited metabolic routes

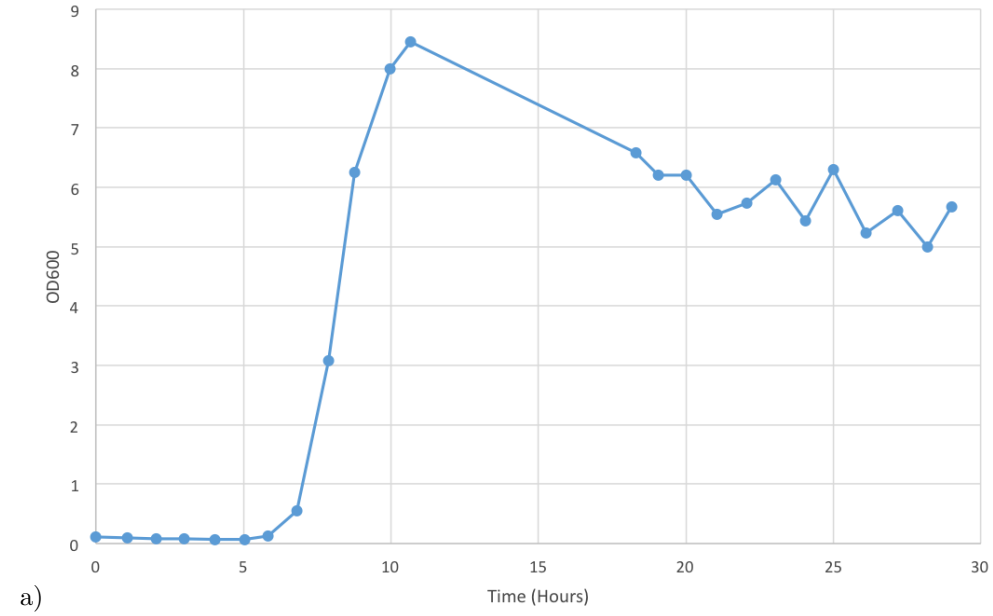
6.1 Introduction

Although *Geobacillus thermoglucosidasius* NCIMB 11955 is described as a facultative anaerobic bacterium, in laboratory conditions it has been found that under true anaerobic conditions the microorganism's growth is significantly impaired (see Figure 6.1).

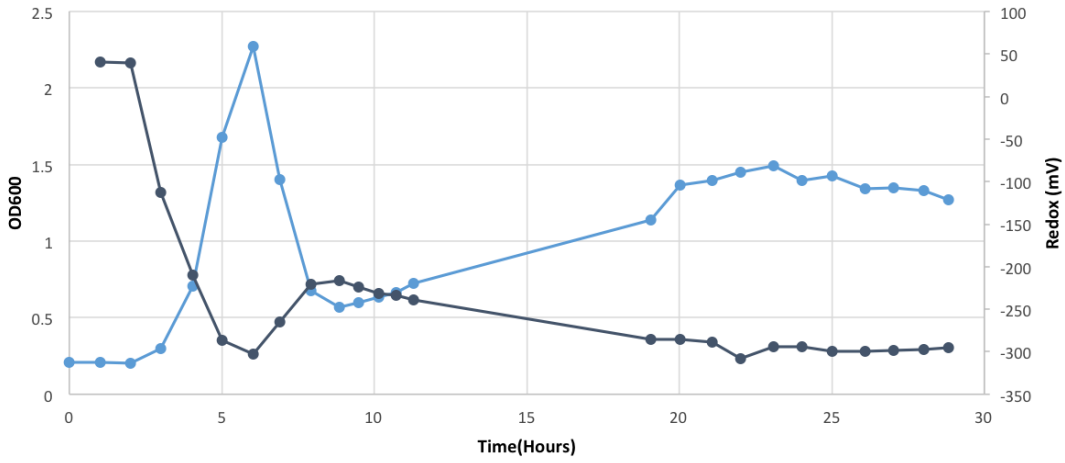
When oxygen is scarce, bacterial growth must be maintained by generation of ATP. This can be achieved either through scavenging for traces of oxygen using high-affinity terminal oxidases or by using alternative final electron acceptors, such as nitrate. In this chapter an attempt to restore growth under anaerobic conditions is described with relation to three highlighted pathways by genome-scale metabolic model: nicotinate and nicotinamide metabolism, thiamine metabolism and porphyrin metabolism. A list of reactions and pathways was made which require oxygen (Figure 6.2). From that list the reactions were selected based on their usefulness in vitamin biosynthesis or cofactor production (reactions which would cause a bottleneck in the pathway and could not be substituted by in any other non-oxygen requiring reaction). The complete list can be found in supplementary materials, the ones investigated *in vivo* (bioreactors) are presented in this section.

The preliminary transcriptional data (RNA-seq) under aerobic and anaerobic conditions was available for *Geobacillus thermoglucosidasius* NCIMB 11955. The data at the time of writing did not have a complete set of biological and technical triplicates, however, the up- and down-regulation of certain pathways was used as a reference point whilst choosing oxygen-limited routes. The RNA-seq data was generated by Dr Shyam Masakapacli and Dr Leann Bacon. The author did the analysis of the preliminary RNA-seq data. CLC Genomics was used to

map the RNA reads to the genome annotations. The gene expression was analysed using *DESeq2* [3] according to protocol by Luo, 2014 [77], combined with *GAGE* tool [79] workflow in R/Bioconductor environment. The RNA-seq data was repeated in triplicates under both aerobic and anaerobic conditions. This data was used in this chapter to see if the genes highlighted as by metabolic predictions as "oxygen limited" steps are indeed up-regulated genes in the anaerobic conditions.



a)



b)

Figure 6.1: The profile of a growth curve for *Geobacillus thermoglucosidasius* NCIMB 11955 in aerobic (a) and anaerobic (b) conditions (The dark grey line shows the redox). The cells were grown in ASM medium with 0.1% yeast extract and 0.1 % tryptone. The oxygen was stopped when the OD reached 1.0 (after approximately 3.5 hours). Courtesy of Alice Marriott.

Figure 6.2: The oxygen limited pathways as highlighted by metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955. Yellow, violet, blue and red denote the degree of interest in a given reaction, with yellow denoting most promising reaction and red, the route least likely to be limiting the growth of *G. thermoglucosidasius* NCIMB 11955 under anaerobic conditions.

6.2 Nicotinate and nicotinamide metabolism

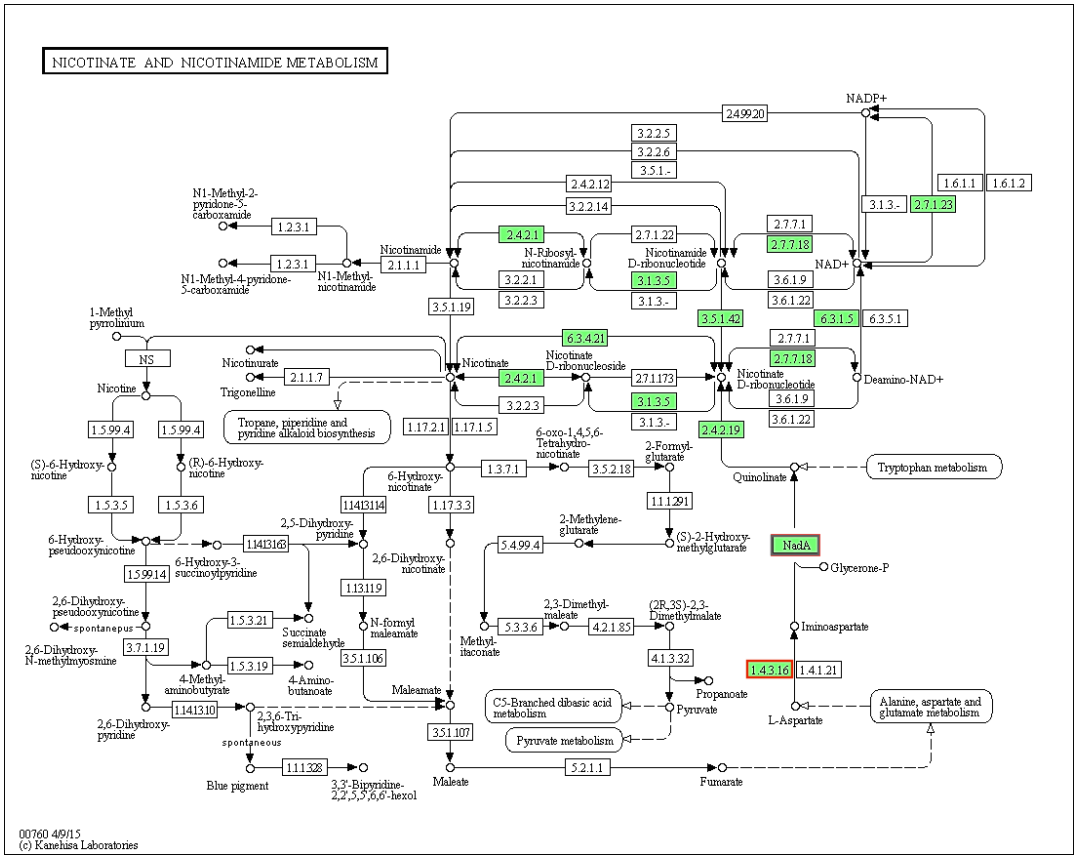


Figure 6.3: Nicotinate pathway with oxygen limited step (red).

NAD is an essential molecule, which plays a major role in redox biochemistry, energy metabolism and has a signalling role in a variety of cellular processes [19]. In a pathway leading to *de novo* NAD^+ biosynthesis, one of the first steps is limited by oxygen (see Figure 6.3 (EC 1.4.3.16)). The reaction is facilitated by L-aspartate oxidase (NadB) and converts L-aspartate to iminosuccinate according to the following reaction :

$$L - aspartate + O_2 \rightarrow iminosuccinate + H_2O_2 \tag{6.1}$$

L-aspartate oxidase from *E. coli* can utilise both oxygen and fumarate as electron acceptor, with the latter yielding succinate. The advantage of substituting oxygen with fumarate makes the reaction possible also under anaerobic conditions [19]. In *E. coli* the enzyme has been found to be actually more efficient with fumarate as an electron acceptor than with oxygen. The enzymatic investigation of NadB in *Bacillus subtilis* has suggested that even though fumarate can be used as a substrate in anaerobic conditions as it is the case in *E coli* the

substrate inhibition was much greater in *Bacillus subtilis* than that in *E. coli* [83].

This enzyme however has not been characterised in any of the members of *Geobacillus* spp., which would elucidate the thermostability and proven kinetic efficiency of L-aspartate oxidase in anaerobic conditions.

The BLAST searches on this enzyme show a low sequence similarity with both mesophilic *Bacillus* spp. and *E. coli* (see Figure 6.4). Indeed, the BLAST analysis reveals that the *Geobacillus thermoglucosidasius* shares less than 70% sequence identity with strains from other *Geobacillus* spp. and shares 35% query cover with 64% of sequence identity with a *Bacillus* sp. X1.

6.2.1 Bioreactor

The L-aspartate oxidase is a crucial step in biosynthesis of NAD. Oxygen is a crucial component of the enzymatic reaction leading from L-aspartate to iminoaspartate. Since this reaction is oxygen-dependent so it can be hence assumed that under anaerobic conditions, *G. thermoglucosidasius* NCIMB 11955 cannot produce efficiently NAD through this route. To test if this is true 0.1% nicotinic acid was added to the chemostat culture after the oxygen was switched off 6.6.

The results do not show restoration of the dwarfed growth but instead mimic the control. The maximum growth rates under anaerobic conditions were: 0.72 h^{-1} with the addition of nicotinic acid and 0.42 h^{-1} without (control). In Micro-aerobic conditions, the recorded growth was: 0.120 h^{-1} for control and 0.114 h^{-1} . The results show that addition of nicotinic acid had little or no effect on cells growing under either set of conditions, with the cell density increasing linearly in all cases, consistent with growth being oxygen limited. The effect of niacin on anaerobic growth hence remain inconclusive.

Description	Max score	Total score	Query cover	E value	Ident	Accession
Geobacillus thermoglucosidasius C56-YS93, complete genome	2825	2825	100%	0.0	100%	CP002835.1
Geobacillus sp. Y4.1MC1, complete genome	2792	2792	100%	0.0	99%	CP002293.1
Geobacillus sp. WCH70, complete genome	1700	1700	98%	0.0	84%	CP001638.1
Geobacillus sp. LC300, complete genome	434	434	70%	3e-117	69%	CP008903.1
Geobacillus sp. 12AMOR1, complete genome	430	430	70%	4e-116	69%	CP011832.1
Geobacillus sp. JF8, complete genome	423	469	76%	6e-114	69%	CP006254.2
Geobacillus thermodenitrificans NG80-2, complete genome	408	408	72%	1e-109	69%	CP000557.1
Geobacillus sp. C56-T3, complete genome	289	289	71%	9e-74	66%	CP002050.1
Geobacillus stearothermophilus 10, complete genome	286	286	71%	1e-72	66%	CP008934.1
Geobacillus sp. GHH01, complete genome	286	286	71%	1e-72	66%	CP004008.1
Geobacillus sp. Y412MC52, complete genome	286	286	71%	1e-72	66%	CP002442.1
Geobacillus sp. Y412MC61, complete genome	286	286	71%	1e-72	66%	CP001794.1
Geobacillus kaustophilus HTA426 DNA, complete genome	286	286	71%	1e-72	66%	BA000043.1
Geobacillus thermoleovorans CCB_US3_UF5, complete genome	277	277	70%	6e-70	66%	CP003125.1
Anoxybacillus gonensis strain G2, complete genome	262	312	83%	1e-65	66%	CP012152.1
Anoxybacillus flavithermus WK1, complete genome	246	246	70%	9e-61	66%	CP000922.1
Jeotgalibacillus sp. D5, complete genome	107	107	36%	6e-19	65%	CP009416.1
Acidobacterium capsulatum ATCC 51196, complete genome	96.9	96.9	6%	1e-15	82%	CP001472.1
Azoarcus sp. KH32C DNA, complete genome	86.0	86.0	5%	2e-12	83%	AP012304.1
Bacillus sp. X1(2014), complete genome	82.4	82.4	34%	2e-11	64%	CP008855.1

a)

Description	Max score	Total score	Query cover	E value	Ident	Accession
L-aspartate oxidase [Geobacillus thermoglucosidasius]	1040	1040	99%	0.0	100%	WP_041270042.1
L-aspartate oxidase [Geobacillus thermoglucosidasius C56-YS93]	1037	1037	99%	0.0	100%	AEH47037.1
MULTISPECIES: L-aspartate oxidase [Geobacillus]	1029	1029	99%	0.0	99%	WP_003248909.1
L-aspartate oxidase [Geobacillus sp. Y4.1MC1]	1028	1028	99%	0.0	99%	ADP73783.1
L-aspartate oxidase [Geobacillus sp. WCH70]	848	848	98%	0.0	83%	WP_015864639.1
L-aspartate oxidase [Geobacillus caldoxylosilyticus NBRC 107762]	800	800	99%	0.0	77%	GAJ38216.1
L-aspartate oxidase [Geobacillus caldoxylosilyticus]	798	798	98%	0.0	77%	WP_033313772.1
L-aspartate oxidase [Geobacillus stearothermophilus]	769	769	98%	0.0	74%	WP_043905610.1
L-aspartate oxidase [Bacillus alveayuensis]	655	655	99%	0.0	63%	WP_044894660.1
L-aspartate oxidase [Anoxybacillus tepidamans]	634	634	98%	0.0	63%	WP_027407946.1
L-aspartate oxidase [Anoxybacillus sp. ATCC BAA-2555]	633	633	98%	0.0	63%	WP_044744670.1
L-aspartate oxidase [Geobacillus thermodenitrificans]	632	632	98%	0.0	62%	WP_029760658.1
L-aspartate oxidase [Geobacillus thermodenitrificans]	631	631	98%	0.0	62%	WP_011887899.1
L-aspartate oxidase [Geobacillus sp. G11MC16]	626	626	98%	0.0	62%	WP_008881058.1
MULTISPECIES: L-aspartate oxidase [Bacillaceae]	625	625	98%	0.0	63%	WP_044744808.1
MULTISPECIES: L-aspartate oxidase [Geobacillus]	611	611	98%	0.0	62%	WP_011232077.1
MULTISPECIES: L-aspartate oxidase [Geobacillus]	611	611	98%	0.0	62%	WP_014196416.1
L-aspartate oxidase [Geobacillus stearothermophilus]	610	610	98%	0.0	62%	WP_049624751.1

b)

Figure 6.4: BLAST analysis for nucleotide(a) and translated nucleotide(b) sequence identity for L-aspartate oxidase.

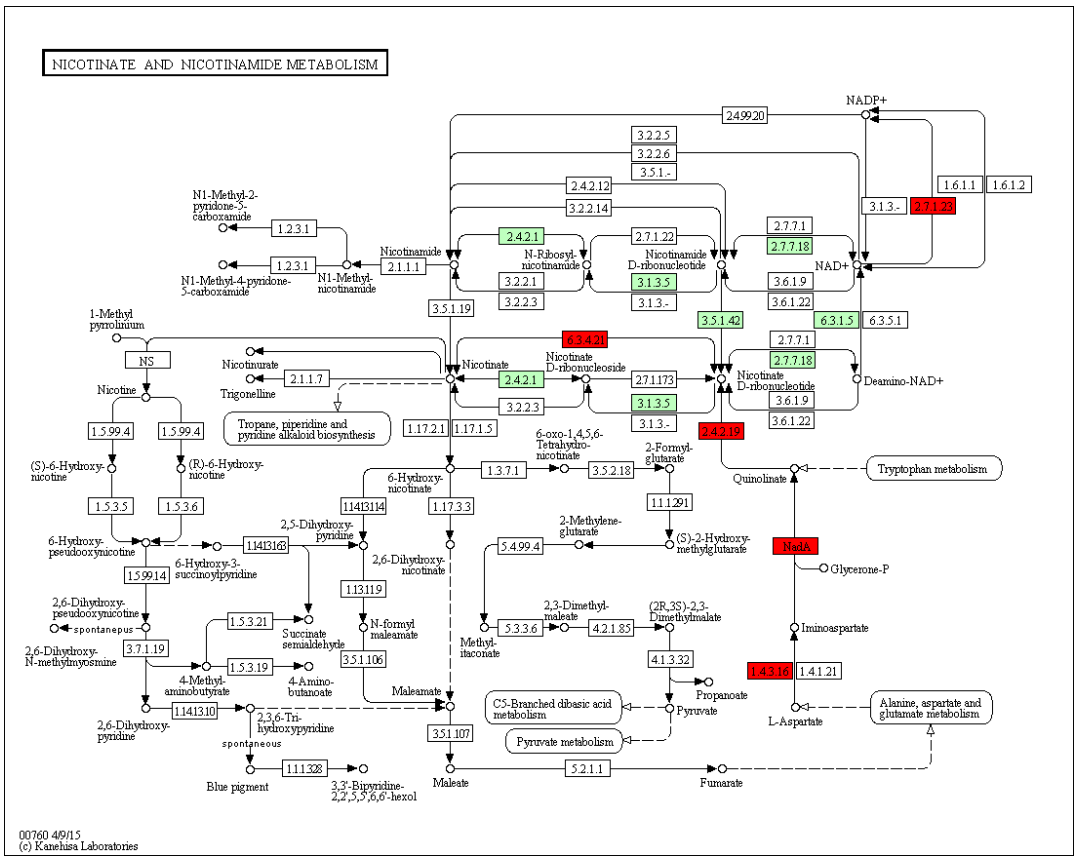
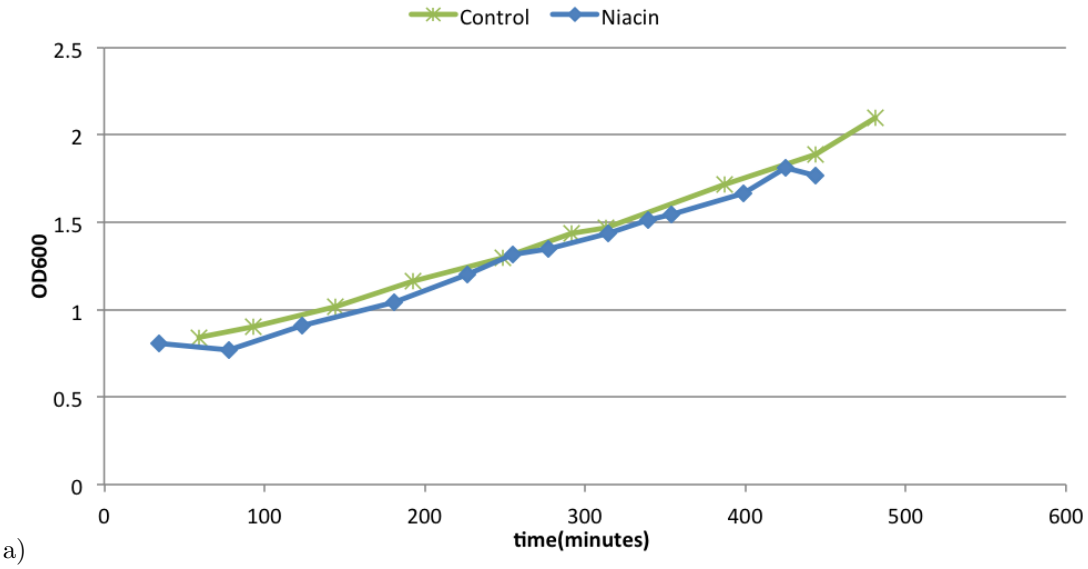
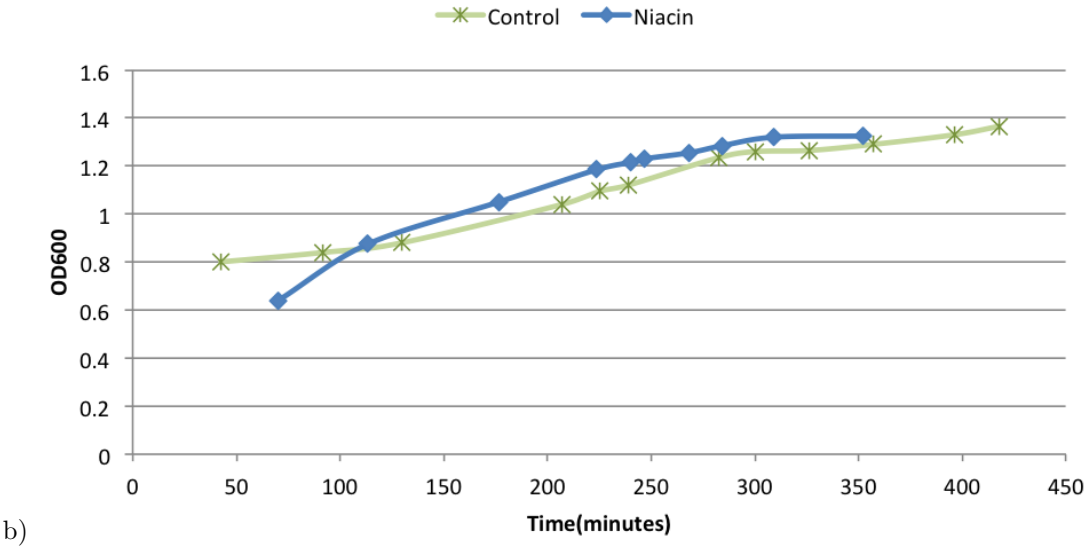


Figure 6.5: Nicotinate and nicotinamide metabolism pathway up-regulated genes (red) under oxygen limited conditions, as suggested by RNA-Seq data. The picture shows how gene regulation is affected in the anaerobic conditions. The L-aspartate oxidase is the first step in the NAD biosynthesis. It is also the only step in this pathway that requires oxygen. The first three enzymatic reactions leading from L-aspartate to nicotinate D-ribonucleotide are hence up-regulated to balance for the decreased efficiency of L-aspartate oxidase as caused by the lack of oxygen.



a)



b)

Figure 6.6: *Geobacillus thermoglucosidasius* NCIMB 11955 growth in chemostat with ASM media (without yeast extract) enriched with 0.1% nicotinic acid (blue) and without (green) in microaerobic (a) and anaerobic conditions (b). Microaerobic conditions are defined at a rate of 50 ml/min of air at 200 rpm with Redox -200 to -230 mV. Anaerobic conditions were achieved by stopping the air flow, when the cell OD reached 1.0.

6.3 Thiamine metabolism

Glycine oxidase(EC 1.4.3.19) is an important homotetrameric flavoenzyme that can facilitate oxidative deamination of a myriad of amino acids [87]. This enzyme catalyses a conversion of glycine to iminoglycine in the biosynthetic pathway for thiamine (Figure 6.7). This enzyme has been characterised for *Bacillus subtilis* and can act on glycine, sarcosine, N-ethylglycine, D-alanine, D-alpha-aminobutyrate, D-proline, D-pipecolate and N-methyl-D-alanine [87]. This flavoenzyme with non-covalently bound FAD [87] has been highlighted in the analysis of oxygen limiting pathway due to its bottleneck effect on production of iminoglycine, which requires oxygen as electron acceptor and no anaerobic substitutes have been identified to date. In anaerobic microorganisms such as *Clostridium* spp. this step is substituted with 2-iminoacetate synthase (EC 4.1.99.19), which converts L-tyrosine to iminoglycine without the need for oxygen.

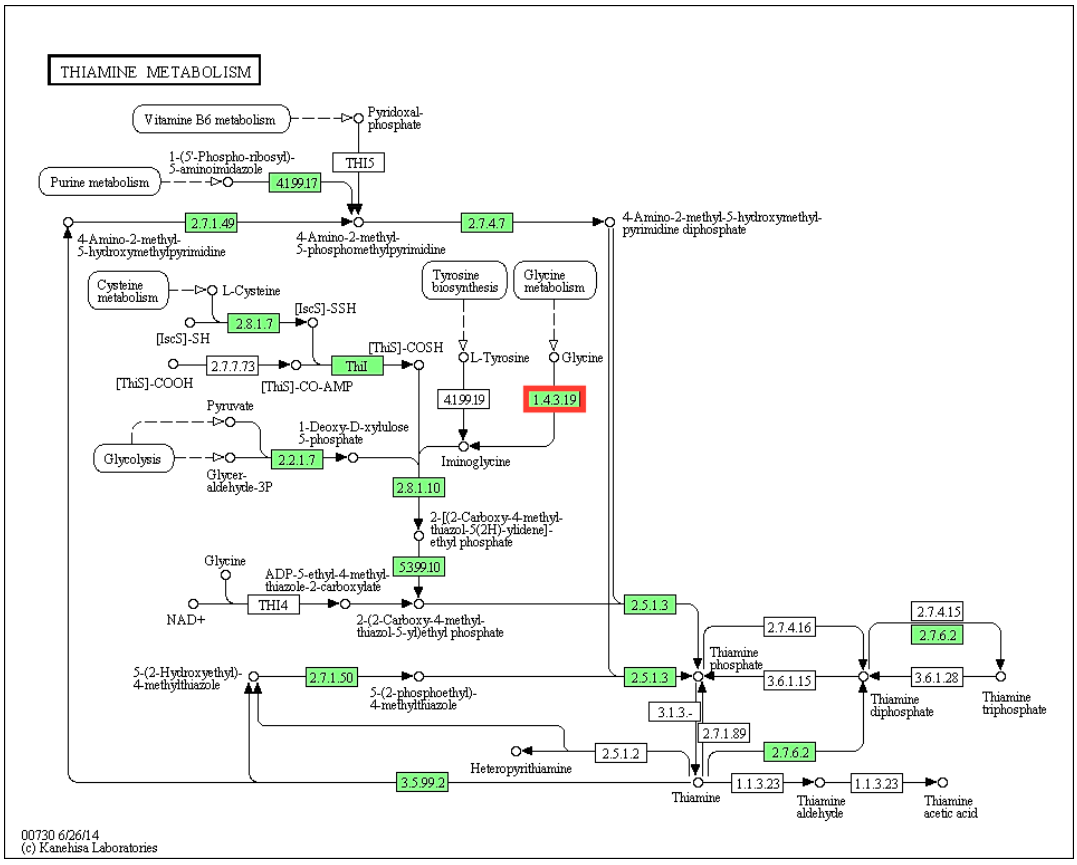


Figure 6.7: Thiamine pathway annotated for *Geobacillus thermoglucosidasius* with oxygen limiting step (red)

6.3.1 Bioreactor

The chemostat cultures were grown in rich medium (ASM medium without yeast extract) with and without addition of thiamine (Figure 6.8). Thiamine was added at the same time as when the air was stopped in the anaerobic conditions (Figure 6.8 a). The result suggest that the addition of thiamine does not restore growth completely, which suggests that the problem of arrested growth during anaerobic conditions can be a more complex problem. With minimal air addition (50 ml/min) 0.174 h^{-1} the chemostat cultures are overall showing a higher growth rate than the control cultures.The growth rate before thiamine addition was 0.168 h^{-1} and following thiamine addition increased to 0.222 h^{-1}. The growth rate of 0.168 h^{-1}(before thiamine addition) is very similar to the growth rate of the control. The increase in growth rate after the addition of thiamine indicates that the thiamine is having a positive effect on the growth. As shown in the Figure 6.7, there is only one route of thiamine biosynthesis in *G. thermoglucosidasius* NCIMB 11955 that requires iminoglycine as a substrate for a reaction catalysed by the thiazole synthase (EC 2.8.1.10). The iminoglycine can by either produced from L-tyrosine or from glycine. Unfortunately, the *G. thermoglucosidasius* NCIMB 11955 does not have a gene annotated for the 2-iminoacetate synthase that converts the L-tyrosine to iminoglycine. The iminoacetate is predominantly found in the genomes of strict anaerobic bacteria such as *Clostridium* spp. since it does not utilise oxygen, unlike glycine oxidase, to produce iminoglycine. If glycine oxidase was the key to unlocking the limited growth of *G. thermoglucosidasius* NCIMB 11955, the addition of thiamine (the end-product of the pathway) should result in improved growth profile of this bacterium under anaerobic conditions.

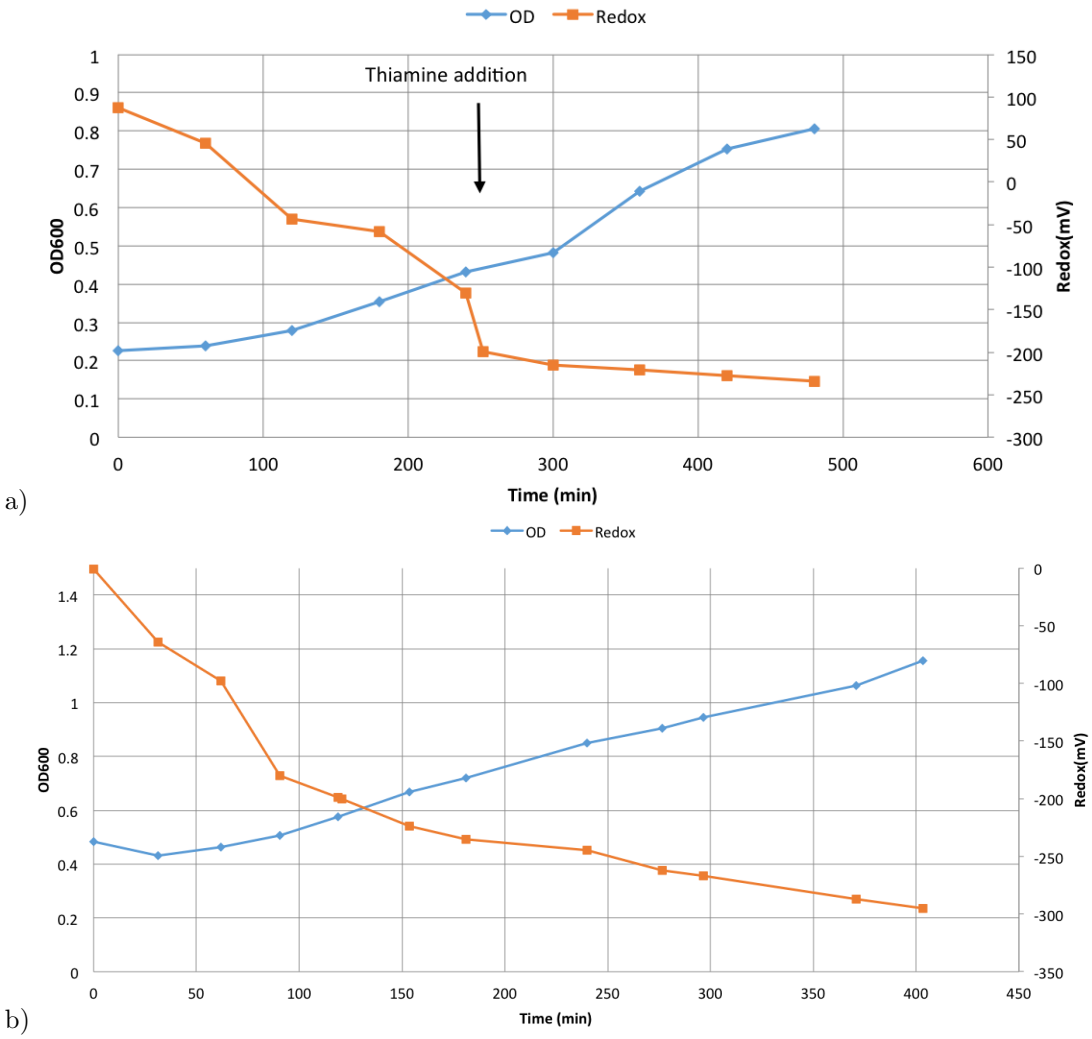


Figure 6.8: Two growth profiles of *Geobacillus thermoglucosidasius* NCIMB 11955 growth in chemostat with ASM media (without yeast extract) enriched with 0.1% thiamine (blue) under anaerobic (a). In anaerobic conditions thiamine was added when the redox value reached -200 mV. The orange line represents the Redox change. Plot b represents growth profile of *Geobacillus thermoglucosidasius* NCIMB 11955 under microaerobic conditions (50 ml/min of air) in ASM medium (without yeast extract) with 0.1% thiamine

6.4 Porphyrin metabolism

Heme is an important compound in the metabolism, playing an important role in both metabolism ad as a regulatory molecule. It has been found to be crucial in the regulation of protein expression on both transcriptional and translational levels. Heme is an important factor in redox-linked reactions such involving oxidases, catalyses or electron transport chains. Heme has been found to increase bacterial growth but can also be toxic to an organism if supplemented in too high levels in laboratory conditions [43]. Protoporphyrinogen IX oxidase (EC 1.3.3.4, HemY) is one of the main enzymes responsible for production of protoporphyrin IX and as is a part of penultimate steps in pathway leading to biosynthesis of protoheme IX (see Figure 6.9) [57].

The HemY catalyses the following reaction:

$$protoporphyrinogen \quad IX + 3O_2 \rightarrow protoporphyrin \quad IX + 3H_2O_2 \tag{6.2}$$

In species capable of efficient growth under anaerobic conditions (such as *E.coli* or *Clostridium* spp.), this enzyme is instead substituted with oxygen independent HemY (EC 1.3.5.3). In this instance, the protoporphyrin IX is produced with menaquinone as electron acceptor, instead of oxygen.

HemY mutants in *Bacillus subtilis* cannot grow without the addition of hemin [57], making it a limiting step in the production of protoheme [57]. HemY from *Bacillus subtilis* can oxidise coproporphyrinogen III to coproporphyrin and protoporphyrinogen IX to protoporphyrin [57].

6.4.1 Bioreactor

Geobacillus thermoglucosidasius NCIMB 11955 was grown on ASM media supplemented with hemin under microaerobic conditions (50 ml/min at 200 rpm). The haemin (15 µM) was added at when the Redox reached 212 mV after 3 hours (Figure 6.11). It was however observed that addition of haemin at this concentration was immediately followed by cell death. Since it has been reported in the literature that haemin if added in a too high concentration, can have a toxic effect [4] on cell growth, a test of *Geobacillus thermoglucosidasius* NCIMB 11955 tolerance to haemin was conducted as described in the next section.

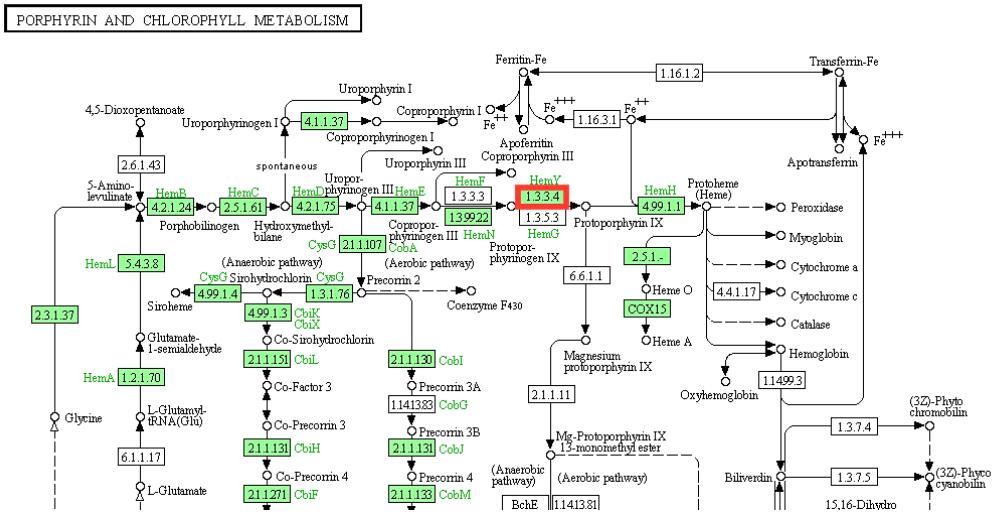


Figure 6.9: Porphyrin metabolism pathway with oxygen limiting step (red).

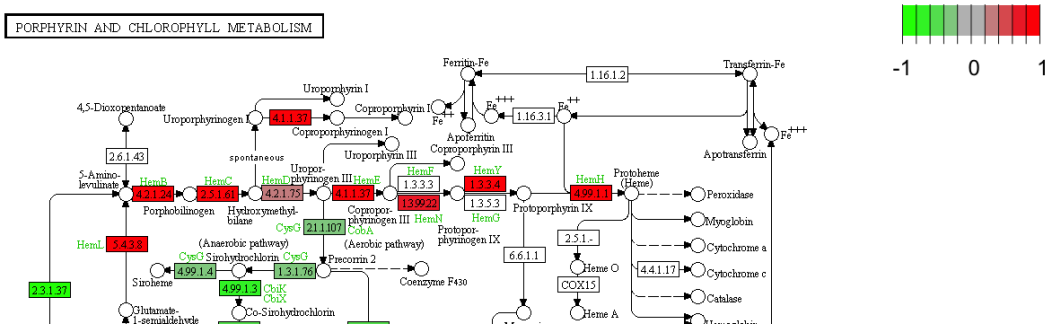


Figure 6.10: Porphyrin metabolism pathway with up-regulated (red) genes and down-regulated genes (green), as denoted by the RNA-Seqdata.

6.4.2 Determination of haemin tolerance levels

To address the problem of haemin toxicity on *G. thermoglucosidasius* NCIMB 11955, two sets of experiments were carried out with the following aims:

- To determine whether G. thermoglucosidasius can grow in medium containing hemin and what level becomes toxic to growth.
- To determine whether the addition of hemin changes the growth of *G. thermoglucosidasius* in culture under microaerophilic conditions.

It was firstly determined that the control culture (0 µM hemin) grew well to an OD of 2.5 after 6.41 hours. In comparison, all of the cultures containing hemin suffered a long lag phase with growth occurring between 6.41 and 22 hours. It is not possible to say whether the cultures had followed a similar growth trend to the control and peaked overnight and declined

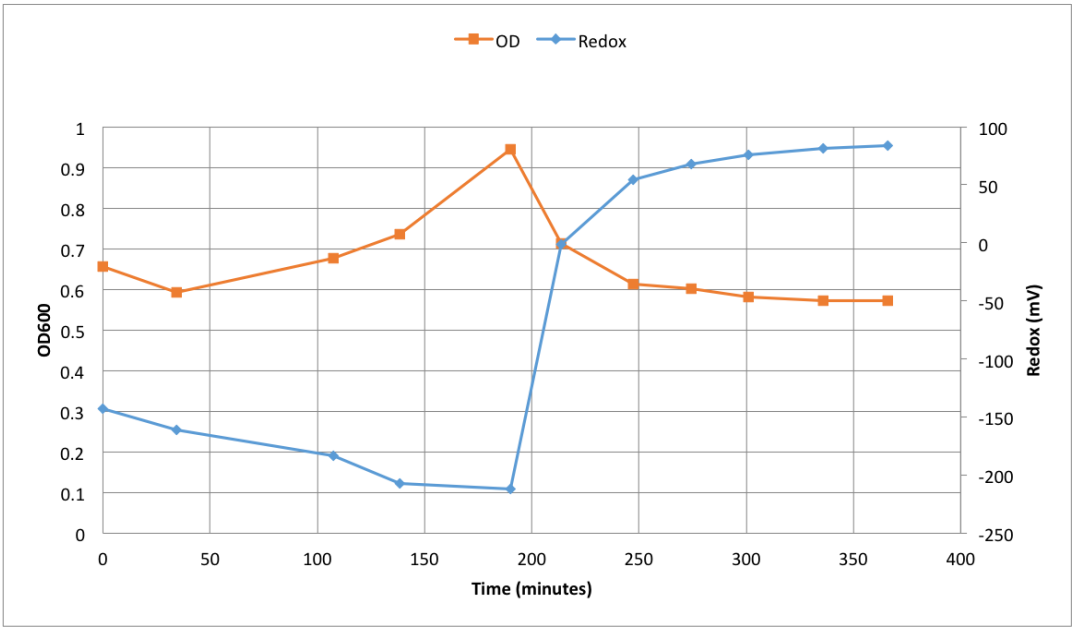
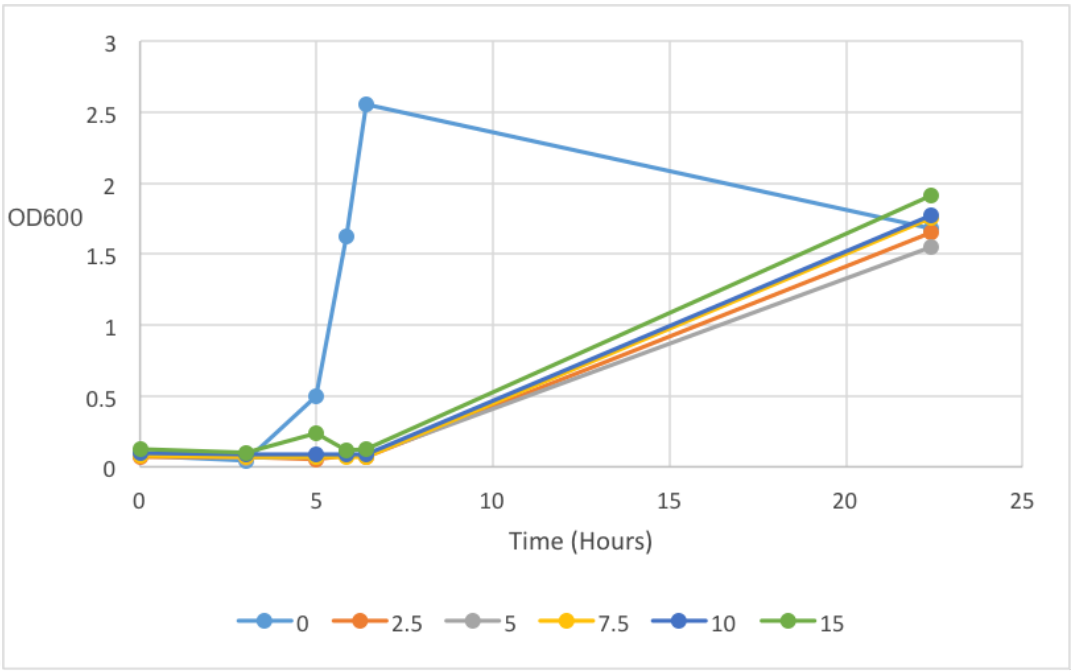


Figure 6.11: Growth profiles of *Geobacillus thermoglucosidasius* NCIMB 11955 growth in chemostat with media enriched with 15 µM of haemin. The OD is denoted by orange line and redox by blue. Haemin was added when the Redox values reached 212 mV.

in OD to the point which was measured at 22 hours. The alternative is that the cultures had a reduced growth rate and grew slowly overnight (see Figure 6.12). It was also noted that there were clumps present in the hemin cultures containing 7.5, 10 and 15 µM hemin and this was recorded at 5 hours after inoculation through to the end of the experiment. This experiment could be extended to determine whether the cultures were still growing or whether they had reached their peak overnight and were in death phase.

The results for the second experiment (see Figure 6.12 b) show that These results show that all cultures grew well. The ODs decreased after hemin addition (6.5 hours) but unfortunately the ODs were already declining at this point. It was noted once again that there was clumping and foaming in the cultures with the highest concentrations of hemin added. The amount of decline was very similar in all cultures which suggests any effect of hemin at these concentrations is concentration independent.

a)



b)

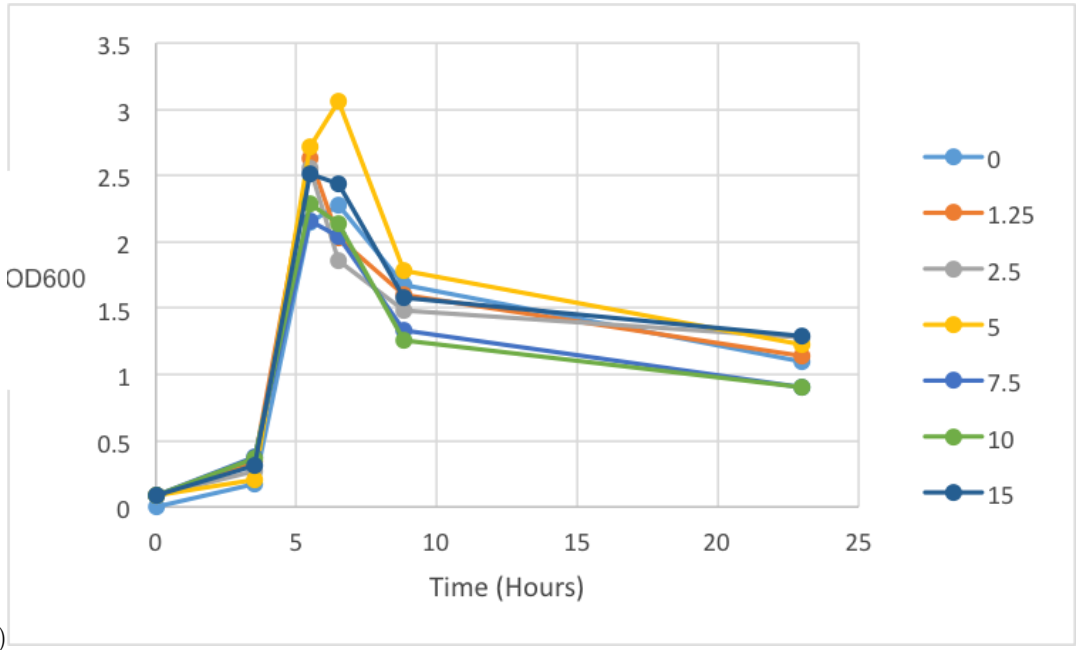


Figure 6.12: Two sets of experiments designed to understand the toxic effect of haemin on *Geobacillus thermoglucosidasius*. Plot (a) show growth profiles in flasks with the haemin concentration ranging from 0 to 15 µM. Plot (a) shows growth in aerobic conditions and plot b) in microaerobic environment . In growth profiles in plot (b) haemin at a range of concentrations was added 6.5 hours after inoculation. The cells were grown in falcon tubes in ASM rich media (with 0.1 % of yeast extract) .

6.5 Conclusions

The results do not give a definitive answer as to the reasons for an impaired growth of *Geobacillus thermoglucosidasius* NCIMB 11955. The hypothesis relating haemin has proved inconclusive and toxicity of the compound could be to blame. Following the set of experiments, described in the previous section, the next steps would be to:

- Repeat the experiments with longer timecourse/ more timepoints (Figure 6.12 (a)).
- Repeat the experiment with hemin addition earlier than 6.5 hours in the growth curve (Figure 6.12 (b)).
- Perform these experiments under aerobic conditions to see whether the growth profile changes depending on aeration.
- Choose conditions and run in the bioreactor.

Although, the niacin addition did minimally improve growth under microaerobic and anaerobic conditions, the full growth profile could not be restored through the attempts described In this chapter. It is possible that the impaired growth is a result of oxidative protein folding pathways. This pathway has been predominantly studied in gram-negative *E. coli* (process catalysed by Dsb family of oxidoreductases), however homologues genes have been found in the genome of *Bacillus subtilis* [97]. This suggests that gram-positive bacteria might have similar mechanism of disulfide-bond formation as that found in their gram-negative counterparts [97]. This would suggest why the metabolic model predicts bacterial growth with the fermentation products, whilst the growth is impaired under truly anaerobic conditions.

Chapter 7

General Conclusions and Future Work

Figure 7.1: Superimposed picture of the *Geobacillus thermoglucosidasius* NCIMB 11955 with its genome-scale metabolic map.

"Essentially, all models are false but some are useful" George E.P. Box.

The idea for this research originated in the advent of metabolic models, when the academic community embraced the possibilities that a genome scale view of metabolism could bring. Four years later, GEMs, having proved their usefulness in generating model-driven genetic modification for enhanced production of a specific product in microbial systems or in providing insight into the physiology of the modelled organism [147, 102, 48, 75]. However, the genome-scale metabolic models are in their core representations of metabolic reactions as deduced from genome annotations but they cannot quantitatively show the behaviour of those reactions within different conditions [35], nor, as argued by Zhang, 2015 [147], can they accurately account for genetic interactions. In the context of *in silico* strategies aimed at metabolic engineering, targeting reactions rather than genes, GEMs simulations do not provide a comprehensive view of the role of isoenzymes or protein complexes and make some predicted strategies infeasible in *in vivo* conditions [147, 48]. This problem stems from the way gene-protein relations (GPR) are represented in the stoichiometric matrix that does not account for the gene interactions between genes in a given gene set [147]. The integration of information on gene relationships, such as a relationship between a reaction encoded by a given gene and isoenzymes or protein complex, can be challenging and requires correct integration of mathematical expressions [147, 1]. Having those limitations in mind, GEMs provide a unique insight into metabolic capabilities of a given organism. The previous published models were those of the central carbon metabolism [130, 143, 90]. The metabolic model of *Geobacillus thermoglucosidasius* NCIMB 11955, presented in this study, is the first, manually curated model of any species in the genus *Geobacillus*. Furthermore, the model benefits from the experimental data used for the estimation of the biomass.

The aim of generating a high-quality global representation of *Geobacillus thermoglucosidasius* NCIMB 11955 metabolism, has led to the development of *PathwayBooster*, which has allowed for analyses of at least theoretical capabilities of not only *G. thermoglucosidasius* NCIMB 11955 but also the clade "thermoglucosidasius". Indeed, the analysis of the gene assignments and the position of genes with regards to gene operons has led to the study of genome re-arrangements between type-strains of the main clades in this genus. Although, no conclusive evidence could be found as to the reason why the genome re-arrangement has occurred, the core three regions of high conservation have been identified and within one such conserved region a gene cluster that presents a novel way utilised by members of this genus to the biosynthesis of vitamin B_{12}. The use of *PathwayBooster* has highlighted gene conservation in the genus such as the presence of 2-oxoglutarate synthase. The FVA of the

metabolic model of *Geobacilllus thermoglucosidasius* NCIMB 11955 however suggests that the 2-oxoglutarate synthase has the minimum and maximum flux of "0" both in the aerobic and anaerobic conditions. This enzyme plays a crucial role in the reverse Krebs cycle and in the prokaryotic carbon fixation route, along with fumarate reductase and citrate lyase [24]. The annotation of this pathway for the species (Figure 7.2) within the genus suggests that the missing steps are catalysed by the citrate lyase (which is absent in all the strains) and pyruvate:ferredoxin oxidoreductase (uniquely present in *G. thermodenitrificans* CCB-US.3-UF5). This on the other hand begs the question why, citrate lyase is absent in all strains within the genus and cannot be found in the genus *Bacillus*. The next step in the analysis of this route would be to create an *in vivo* 2-oxoglutarate synthase knock-out strain and observe physiological effects of the genetic manipulation.

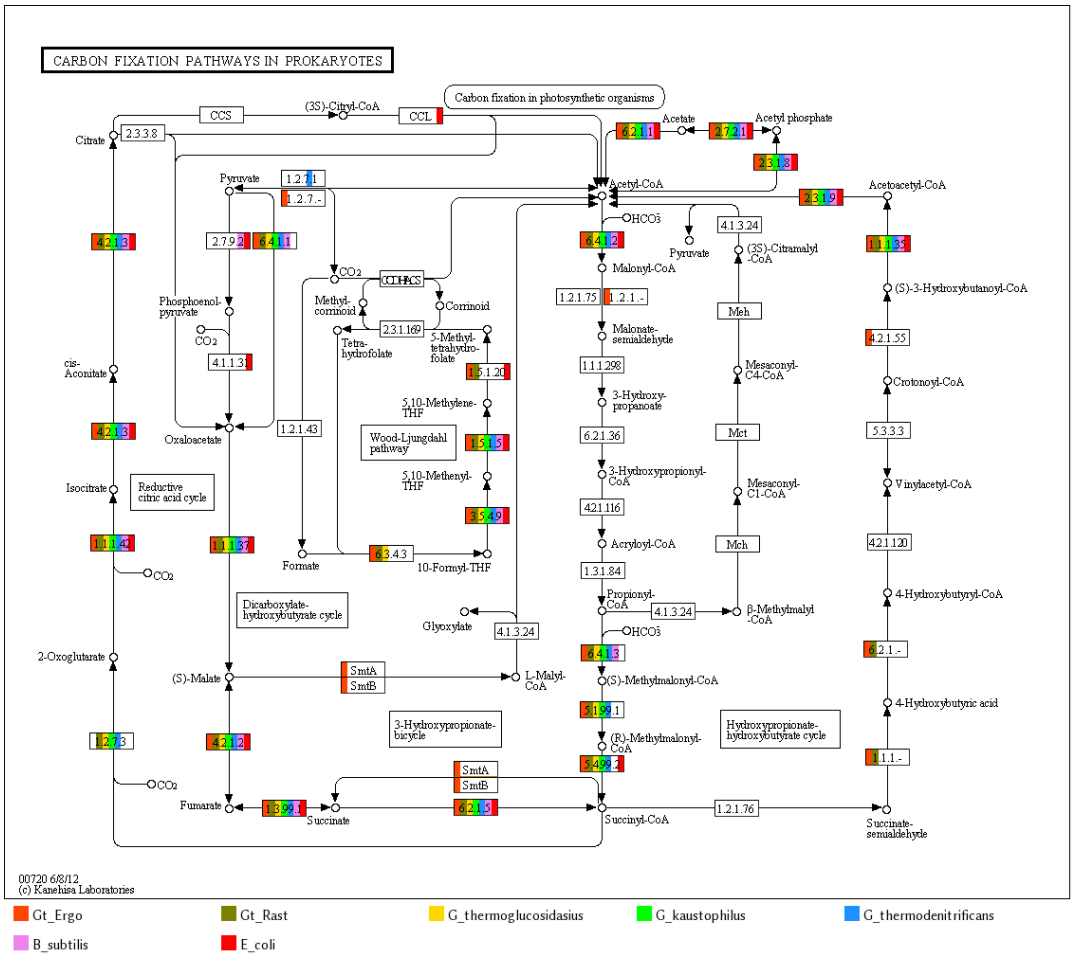


Figure 7.2: Pathway gene annotation for the species within genus *Geobacillus*

One of the unique gene annotations of *Geobacillus thermoglucosidasius* NCIMB 11955 was that of NADP-dependent GAPDH, which has been analysed *in vivo* for the validation of the functional assignment, and brings a novel understanding of the metabolic capabilities of this strain. As described in this thesis, this enzyme may play a crucial role in the generation of NADPH in the absence of the transhydrogenase. It is also interesting to note that the flux through this route has been predicted by the model both in aerobic and anaerobic conditions. The next step in the analysis of this enzyme would be to perform an *in vivo* NADPH-GAPDH gene knock-out to measure the performance of the strain in both aerobic and anaerobic conditions and elucidate the importance of this enzyme for this strain.

The work in this study has highlighted the potential routes for enhanced production of succinate in this strain by double knock-out of pyruvate kinase and malate dehydrogenase. The theoretical product yield of this double mutant was significantly higher than the approach of enhancing an already present route through the addition of glutamate decarboxylase. Although, as discussed earlier the model does not account for the gene relationship or the gene regulation and hence feasibility of this approach needs to be evaluated *in vivo*.

Finally, the focus on genome rearrangements within this genus has highlighted the varied number of transposable elements, which varies depending on the clade. This project has highlighted that the largest number of transposable elements can be found within clade "thermoglucosidasius", which might be a reason why a degree of a genome rearrangement has been also observed.The analysis of the HUS cluster within this clade has highlighted this finding, showing that the genes originally found within this cluster are found in different positions depending on the strain analysed within the clade "thermoglucosidasius".

The next step of curation of this genome-scale metabolic model would include incorporation of "-omics" data, such as transcriptomics, metabolomics or proteomics approach, to get a clear understanding of the genome-scale metabolic network behaviour under different physiological conditions. The incorporation of such data would allow for creation of multiple steady-state models, which would be a representation of the metabolic network under aerobic conditions, microaerobic conditions and anaerobic conditions. Such anaerobic model could be a useful tool for analysing the *Geobacillus thermoglucosidasius* NCIMB 11955 oxygen limitation, which has remained elusive in the analysis discussed in this study.

Chapter 8

Materials and Methods

8.1 Laboratory methods

8.1.1 Bacterial strains and growth conditions

List of bacterial strains and plasmids is shown in table 8.1. *G. thermoglucosidasius* strain 11955 was grown at 60°C at 200 rpm in 2TY medium[1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5%(w/v) NaCl] or ASM medium (minimal media) with the following composition for the latter: 1% (w/v) tryptone, 1%(w/v) yeast extract , 55.5 mM D-glucose, 8.7 mM citric acid , 20.2 mM Mg SO_4, 10mM K_2SO_4, 22.6 mM NaH_2PO_4, 0.8mM $CaCl_2$, 25mM $(NH_4)_2SO_4$ with trace element solution. The rich media is defined as ASM with 0.25% of yeast extract. These were also used as growth medium requirements for the metabolic model.

Composition of trace element solution as follows: 0.5mM $ZnSO_2$, 0.2mM $CoSO_4$, 0.1 mM

Strain or plasmid	Description	Reference
G.thermoglucosidasius NCIMB 11955	*G.thermoglucosidasius* wild type strain	This study
E.coli BL21	Expression strain	Invitrogen, CA
*Bio*Blue	Cloning strain, chemically competent cells	Bioline, UK
pET-28-aa-c(+)	Cloning and expression vector	Novagen, UK

Table 8.1: Bacterial strains and plasmids used in this study. pET-28-aa-c(+) is a cloning and expression vector with kanamycin resistance.

$CuSO_4$, 2.29 mM $FeSO_4$, 0.3 mM $NiSO_4$, 0.9 mM $MnSO_4$, 0.1 mM H_3BO_3, 6mM H_2SO_4.

E. coli strains *Bio*Blue (Bioline) and BL21 were grown at 37℃ in LB broth or LB with supplemented agar and when appropriate, with antibiotic kanamycin (50 µg ml^{-1}).

8.1.2 Sugar Assays

API 50 CH standarised system was used in order establish growth of *G. thermoglucosidasius* NCIMB 11955 on different carbohydrates. . Medium CHB/E was used for this according to product specifications. The test was done for the following carbohydrates: glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose,L-xylose, -D-adonitol, methyl- D-xylopyranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-D-mannopyranoside, methyl-D-glucopyranoside, N-acetylglucosamine, amygdalin, arbutin, ferric citrate, salicin, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-saccharose, D-trehalose, inulin, D-melezitose, D-raffinose, amidon, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol and potassium gluconate. When the cells OD reached 1.0 the aliquots of the bacterial solutions were introduced into the strip. The incubation time for the API 50 was 24 hours at 55℃. Furthermore, cell were grown on maltose, D-mannitol, D-mannose, α,α-trehalose. Similar growth experiments were done on the utilisation of nitrate by the *Geobacillus thermoglucosidasius* NCIMB 11955. The summary of the assays can be found in Table (3.24).

8.1.3 DNA manipulation

Genomic DNA was extracted from *G. thermoglucosidasius* NCIMB 11955 with Wizard SV ® Genomic DNA Purification System (Promega, UK) according to manufacturer's specifications. QIAprep Spin kit (Qiagen, Crawley, UK) was used to prepared plasmid DNA and was used with accordance to manufacturer's instructions. All restriction enzymes used along with T4 DNA ligase were obtained from Life Technologies (Paisley, UK) and were used according to manufacturer's protocols. Eurofins Genomics (Ebersberg, Germany) were used for DNA sequencing. DNA fragments were separated by gel electrophoresis (0.8 % agarose gels) and stained with SYBR® safe (Life Technologies, Paisley, UK).

8.1.4 Preparation of glyceraldehyde-3-phosphate recombinant

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was amplified from *G. thermoglucosidasius* NCIMB 11955 using primers BKL001(F) and BKL002(R) (see Table8.2, Figure3.9).

Amplified fragment was then ligated into pET-28-a-a-(+) vector (Figure 8.1) and transformed into *Bio*Blue chemically competent cells (BioLine, UK) and the correct DNA sequence was confirmed by sequencing. Plasmid containing the ligated gene was then used to transform *E.coli* BL21 for protein expression studies. Transformants were grown overnight in ASM (amended with kanamycin) and allowed to grow to log phase. For determination of optimal expression cells in small scale expression scale (10 ml cultures) were induced with IPTG at different concentrations (0mM, 0.1mM, 0.8 mM) for 3, 5 and 20 hours each and harvested by centrifugation at 13,000 rpm for 5 minutes. Small expression trials were repeated on a large scale (250ml cultures) and transformants were allowed to grow to log phase and harvested by centrifugation(20 minutes, 4000xg) 5 hours post 0.1 mM IPTG induction.

Primer	DNA Sequence	Restriction site
BKL0001(F)	ATTCAA**GCTAGC**AAAGCAAAAGTGGCGATTAATGGGTTT	NheI
BKL0002(R)	TAACAT**CTCGAG**TTAAGCGTTCACACTGATCTTTTCTGCC	XhoI

Table 8.2: Primer sequences.

The His-tagged recombinant protein was purified according to the following protocol. Cell pellet was resuspended in 20mM Tris pH 8.0, 300mM NaCl and 10mM Imidazole (buffer A) at a concentration of 0.4g of cell pellet per 1ml of buffer A. Cells were sonicated on ice for 4 blasts at 10 microns for 30 seconds . Sonicated samples were then centrifuged for 10 minutes at 30,000 x g. The supernatant was collected and stored and the resulting cell pellet debris was discarded. TALON ® metal affinity resin (Clonetech 635502) was used to purify the recombinant protein by a series of imidazole dilutions (1M, 0.5 M, 0.2 M, 0.1M) with 20mM Tris pH8.0 and 300mM NaCl.

8.1.5 SDS-PAGE and size exclusion

Electrophoretical separation of proteins was done using 12% polyacrylamide gels (30% bis-acrylamide, 1M Tris pH 8.8 (pH 6.8 for stacking gel) , 10% SDS, TEMED and 20% APS). Membranes were stained with Coomassie Brilliant Blue (Sigma-Aldrich, Gillingham, UK) and visualised on Gene Box UV camera (Syngene, UK) (Figure 3.11).

Recombinant protein was purified from imidazole buffer through size exclusion chromatog-

pET-28a(+) sequence landmarks	
T7 promoter	370-386
T7 transcription start	369
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites	
(*BamH* I - *Xho* I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	773-1852
pBR322 origin	3286
Kan coding sequence	3995-4807
f1 origin	4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid; subtract 1bp from each site beyond *BamH* I at 198. pET-28c(+) is a 5367bp plasmid; subtract 2bp from each site beyond *BamH* I at 198.

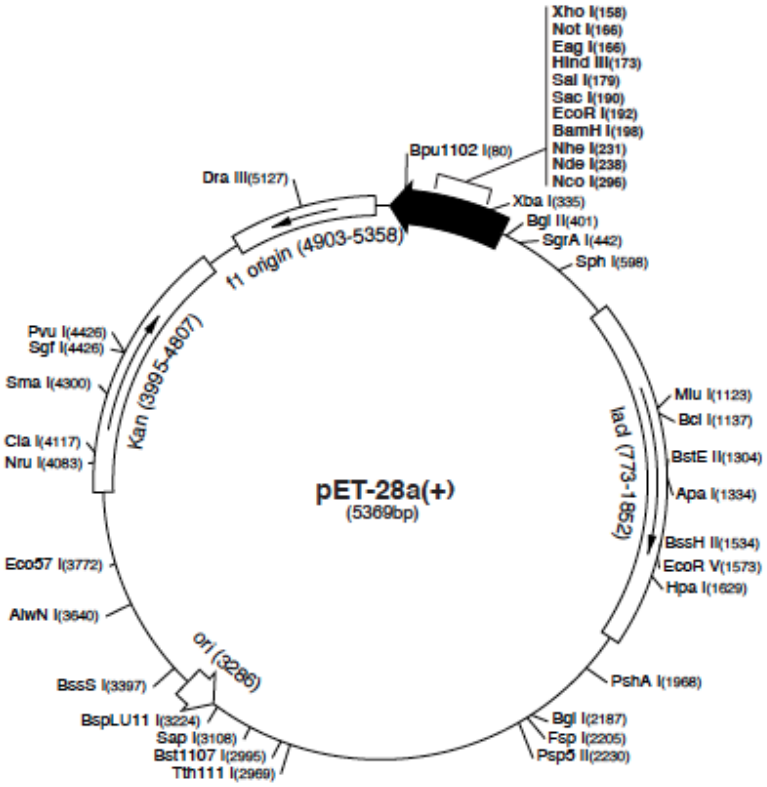


Figure 8.1: pET-28-a expression vector used (Novagen, UK).

raphy using AKTAFPLC system (GE Healthcare, UK) with size exclusion column HiLoad 16/600 Superdex 200 pg (GE Healthcare, UK) and equilibrated with 25 mM His-HCl pH 6.2. GAPDH was eluted (5 mM acetate, 5MM His-HCl, 5mM glutamate, pH 4.0) in 1.0 ml fractions. Protein concentrations were analysed measuring the absorbance at 280 nm and fractions containing highest concentration were pooled together.

8.1.6 GAPDH enzyme assay

Prior to the enzyme assays the protein concentration was estimated through Bradford assay using the protein sample and Bio-Rad Protein Assay Dye Reagent at the 20% (v/v) (Bio-Rad Laboratories GmbH, Munich, German). The standard curve was estimated using bovine serum albumin (Pierce, Rockford, USA) through a series of concentrations ranging from 0-10 µg / ml. Enzyme assays were done according to a modified method of Velick(1955) [142]. The assay was performed in the following reaction buffer: pyrophosphate buffer (0.015M sodium pyrophosphate pH 7.35 with 0.03M sodium arsenate) with 0.1M DTT, either 7.5mM NAD^+ or 7.5mM $NADP^+$ for cofactors pH 7.35. The concentration of GAPDH used in the assay was 0.183 mg/ml in 1ml total of reaction. Reaction buffer was equilibrated to 60℃ in a water bath. The absorbance was measured at 340nm, the change in absorbance was observed after adding 7.5mM D-glyceraldehyde-3-phosphate. The experiments were done in two sets: one for GAPDH activity with NAD^+ and other with $NADP^+$ as a cofactor source. Once the preferred cofactor was identified, the activity of the enzyme was measured. The enzyme kinetics were calculated using SigmaPlot software (London, UK).

8.1.7 Bioreactor

Chemostat culture of *Geobacillus thermoglucosidasius* NCIMB 11955 was conducted by Dr Shyam Masakapalli, Alice Marriott and the author in a 1.5 L fermenter (Applikon, Worcestershire, UK). The conditions chosen were: anaerobic condition, microaerobic conditions and anaerobic conditions.The media used was ASM (as defined in the previous sections), supplemented with haemin (various concentrations), 0.1% yeast extract, or 0.1% thiamine, depending on the experiment conducted at pH 7. The cells were initially grown aerobically until the OD600 reached a value between 3 and 4. The initial culture was used to inoculate the bioreactor. The cells were grown with air flow at 200 rpm. Depending on the conditions, the air supply was either switched to anaerobic condition when the redox reached 200 mV (when the steady state was achieved) or in the case of microaerobic conditions the air rate

was decreased to 50 ml/min. Microaerobic conditions are defined at a rate of 50 ml/min of air at 200 rpm with Redox -200 to -230 mV. Anaerobic conditions were achieved by stopping the air flow, when the cell OD reached 1.0.

The following conditions were used to growth the bacterial cultures in the chemostat prior to biomass composition analysis: temperature = 60 ℃ , pH=6.8, Minimal ASM with 1% Glucose, with total culture volume = 1.5 L in 2 L volume bioreactor. The cultures were grown both in aerobic and anaerobic conditions.

8.2 *In silico* methods

8.2.1 Codon Usage Comparison

The codon usage for type strains in taxa *Geobacillus* was calculated for *G. kaustophilus* HTA426, *G. thermoglucosidasius* C56-YS93, *G. thermoleovorans* CCB US3 UF5, *G. stearothermophilus* NUB3621 and *G. thermodentificans* NG80-2 *B. subtilis* 168. The Percentage of Codon used per amino acid and Frequency codon occurence per 1000 was taken from Codon Usage Database [6]. The codon usage is calculated from NCBI-GenBank Flat File Release 160.0 [June 15 2007].

8.2.2 Genome Annotation and Assembly

ERGO Integrated Genomics ™ was commissioned by TMO Renewables Ltd (now, ReBio Ltd) to sequence, assembly and preliminarily annotate the genome of *G.thermoglucosidasius* 11955 ([104]). To acquire the best possible annotation and minimise misannotations, the genome was also annotated through RAST server. The ERGO and RAST annotations were compared manually, one ORF at a time. The operon structures and continuous genes ware taken into account whilst processing the ORFs. The genes in operons were compared to the operons of the closely related species from genus *Geobacillus* spp. The degree of similarity between annotations were taken into account and were represented by assignment of confidence scores. The confidence scores ranged from 0 to 4. The description of scores can be found in table 8.3.The gene score has been incorporated in the model in the SBML. As the modelling work progressed, the confidence score included the level of knowledge available for a given gene with a view of metabolic model. As such the scores subsequently reflected, from the lowest scores, genes: not evaluated ("0"), included in the model but without literature or biochemical evidence ("1"), annotations stemming from genome annotation or from physiological data ("2"), evidence from genetic focused experimental approach ("3") or including extensive biochemical work on protein expression ("4"). This approach has been adapted from methodology suggested by Thiele, *et al.*, 2010 [134] and Yang *et al.*, 2010 [146].

It should be noted that the confidence score was treated in this study as a label and not as a statistical evidence.

Origin of replication was found using gene order comparison and through closely related gene order comparison. The oriC was set to 1bp and the remaining genes were rearranged

Confidence Score	Description	Example
0	No annotation	EC -.-.-.- /-
1	Annotation found by only one platform	EC 1.1.1.1 / EC -.-.-.-
2	Mismatch in annotation between both platforms	EC 1.1.1.1 / EC 2.2.2.2
3	Difference in depth of annotation	EC 1.1.1.- / EC 1.1.1.1
4	Complete agreement in annotation	EC 1.1.1.1 / EC 1.1.1.1

Table 8.3: Genome annotation confidence scores.

accordingly. RAST platform was also used to generate primary genome arrangement comparison between *G.thermoglucosidasius* 11955 and other closely related species in the genus *Geobacillus*.

8.2.3 Preliminary Model Reconstruction

The data on construction of the metabolic model was done solely by the author. All the scripts and *in silico* analyses done to achieve it was also done by the author. Reconciled annotations were blueprints for preliminary reconstruction of metabolic pathways. Mapping between genes and reactions was done using custom Python scripts that are attached in the supplementary data. Genes with defined enzyme commission numbers are connected to the biochemical database such as KEGG [64, 65] and BRENDA [114, 113] and Entrez Gene [81]. Each reaction formula based on EC annotation was retrieved from KEGG database using KEGG API (application programming interface). The neutral formula of each molecule in a given reaction within the model was retrieved using KEGG API. At this stage manual reconstruction refinement was carried out. The metabolic reactions were manually curated by finding the misassigned or unfeasible metabolic reactions. A tool was created for this purpose with collaboration with Imperial College London called PathwayBooster [74]. PathwayBooster helps finding holes or wrongly assigned reaction by a manner of comparison with evolutionarily-related selected genome . This software employs KEGG pathways for visual representation. In this study the genomes of: *G. kaustophilus* HTA426, *G. thermoglucosidasius* C56-YS93, *G. thermoleovorans* CCB US3 UF5, *G. stearothermophilus* NUB3621 and *G. thermodentificans* NG80-2 *B. subtilis* 168, *G. thermoglucosidasius* ERGO annotation with RAST annotation and *E. coli* for contrast. MEMOSys [95, 94] was used to generate preliminary model in XML format and apply basic thermodynamics rules to the model. However, as explained in previous chapter, a vast amount of manual curation was applied to elucidate

the proper metabolic capabilities of the model. The metabolite nomenclature and identifiers were also generated by MEMOSys platform and as such were kept in the metabolic reconstruction. Metabolites as well as reactions have however also KEGG identifiers which can easily link the model to external databases.

This system allows for managing, developing and storing models. The crude model in the SBML format needed to be created as an input file for the platform. For that purpose scripts were written especially for this project in Python, using libSBML and miniDom from previously described gene annotation. Depending on the models, the XML files were either edited through Python programming language or in COBRA toolbox ([61]). The reaction reversibility was resolved using eQuilibrator [55, 91, 54]: biochemical thermodynamics calculator. This software predicts reaction reversibility based on estimation of thermodynamic parameters (ΔrG and ΔfG) which in turn suggests the directionality of a given biochemical reaction. This was applied to all the reactions present in the model and the upper and lower bounds were set accordingly. If a reaction is a reversible one (reversibility="true") the lower bound was set to -1000 (mmol/gDW/h) and upper bound 1000 (mmol/dGW/h). If the reaction was found to be irreversible (reversibility="false"), the lower bound was set to 0 (mmol/gdw.h) and upper bound to 1000 (mmol/gDW/h). The model was divided into two compartments: extracellular[e] and cytosol [c].

The model versions and updates were tracked using *GitHub* repository. All the reactions with metabolites can be found in the spreadsheet format in appendices.

8.2.4 Manual curation of preliminary metabolic model

Metabolic reconstruction based solely on KEGG database pose problems within the metabolic model. The main problem being introduction of gaps in the metabolic network general assignment. The activity of some enzymes is represented with generic substrates and products such as "fatty acid" or "alcohol". Such instances needed to be substituted with a specific substrate or cofactor to fulfil metabolic model basic assumptions. The replacements were found through manual search of literature or relevant databases such as BRENDA [114, 113]. The complete table with all relevant data sources used in this study can be found in the following table 8.4.

The exchange reactions for all extracellular metabolites were added through custom scripts in Python and miniDom. Spontaneous reactions were added to the reconstruction based on KEGG database. The only spontaneous reactions added were those which metabolites

Name	Description	Stage of reconstruction
GenBank	General nuceleotide sequence database	Gene annotation
SEED	Integrated platform for genome analysis	Gene annotation
BRENDA	Comprehensive enzyme database	Model reconstruction & refinement
UniProt	Universal Protein Resource daatabase	Model reconstruction & refinement
CheBI	Database of small molecules	Model reconstruction
PubChem	Small molecules database	Model reconstruction
KEGG	Database for metabolic pathways with reactions	Model reconstruction & refinement
UniPathway	Database of enzymatic reactions and pathways	Model refinement
Reactome	Database for metabolic pathways	Model refinement
BioCyc	Organism-specific database of pathways	Model refinement

Table 8.4: Databases used for manual curation of the metabolic model of *G. thermoglucosidasius* NCIMB 11955 .

were already present in the preliminary model to limit the number of dead-ends introduced. Preliminary model was analysed for gaps and dead-end metabolites in MATLAB[133] using COBRA toolbox software [61]. The functions used were FastGap and GapFind. Both come back with a list of dead-end metabolites, which are then further classified as root gaps or downstream gaps. These metabolites were investigated with accordance to the graph below. This protocol for gap filling has been modified from the Palsson method [134] (Figure 8.2). This approach of finding gaps is intrinsically connected with Flux Balance Analysis (FBA) which is described in detail in the next sections.

Model visualisations were done using SBMLSchematic and Cytoscape [101, 116] .

The model was also checked for reaction connectivity using *CheckMet* software [73] and the results can be found below 8.3. This software is a machine learning methodology based on simple network topological properties of a model. Reactions are scored from: 0-1 based on their connectivity and fluxes. We have concluded based on those results that the genome-scale metabolic model of *G. thermoglucosidasius* NCIMB 11955 is connected well within its network of reactions.

8.2.5 COBRA toolbox

Flux balance analysis is described in detail in chapter: *metabolism*. FBA was simulated in MATLAB environment using COBRA toolbox (Constraints Based Reconstruction and Anal-

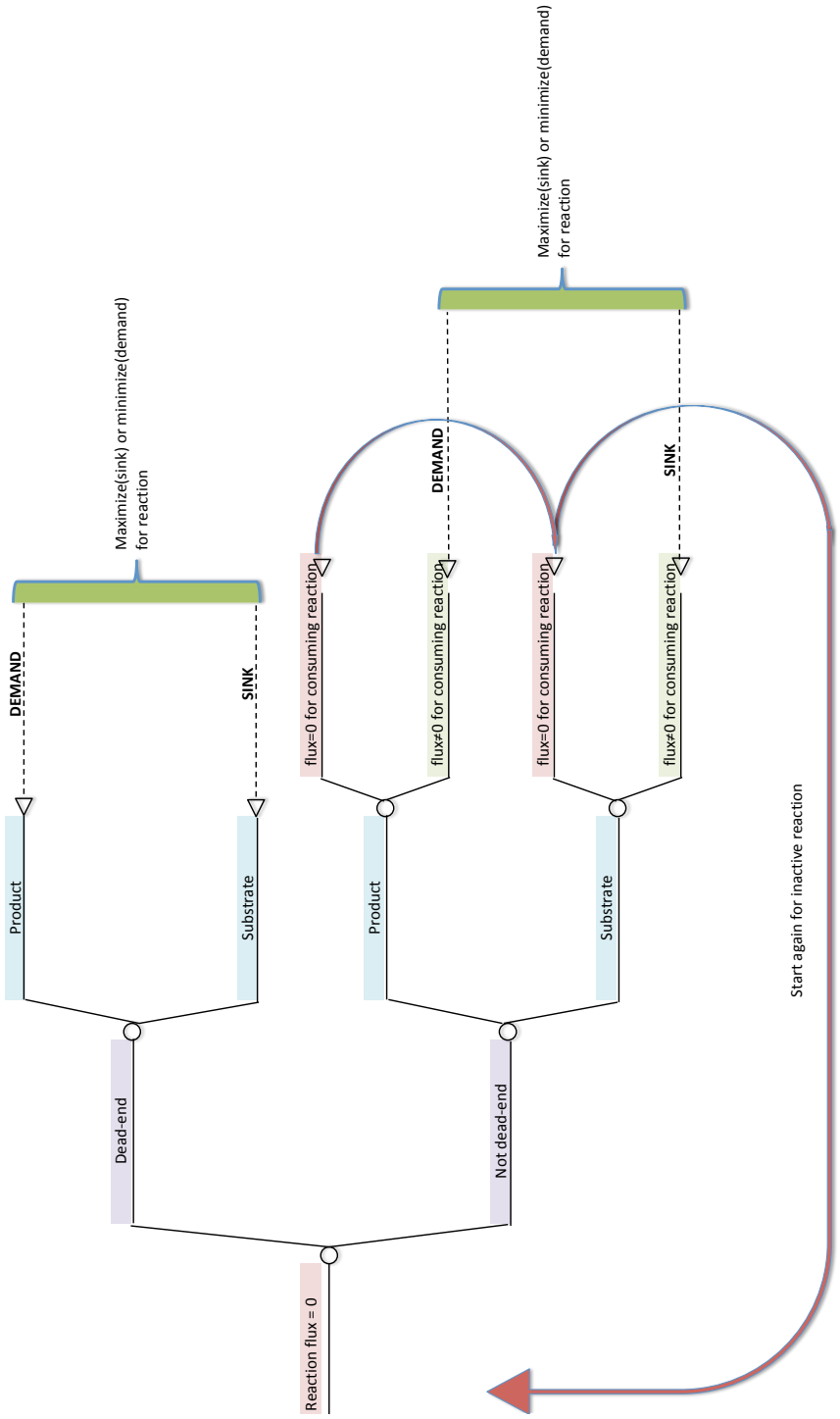


Figure 8.2: Diagram for debugging reactions with fluxes of 0. The above diagram was adopted and modified from the protocol by Thiele *et al*, 2010 [134]

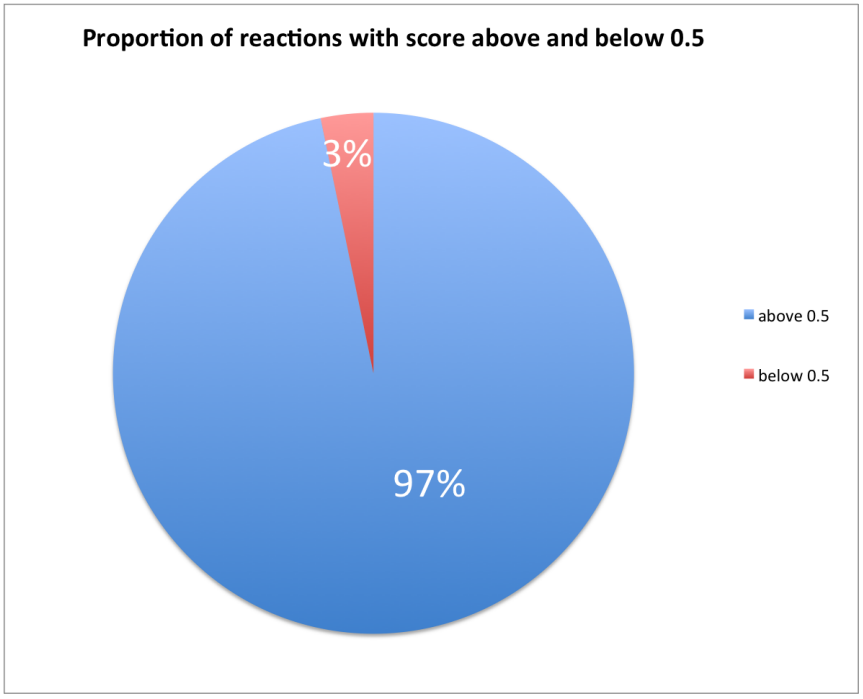


Figure 8.3: Check-met results: proportion of reactions in the model with score over 0.5 and below 0.5.

ysis) [61]. The same toolbox was used for flux variability analysis (FVA) and minimisation of metabolic adjustment (MOMA) approach. All the above mentioned tools required choice of appropriate solver. In this study *IBM* CPLEX and *Gurobi* were chosen. *OptKnock()* was used as a tool for *in silico* knockout approach and subsequent analysis of vector of fluxes which resulted from these permutations. The fluxes for the FVA and FBA for the reactions in the model can be found in the appendices. The model was simulated according to the standard protocols, however the script can be found below for simulation of the FBA under anaerobic conditions.

8.2.6 GAM and non-GAM estimation

Theoretically, the energetic expense for biosynthesis of macromolecules is calculated by firstly establishing what is the percentage of a given macromolecule(protein, DNA, RNA) per dry weight and then the total amount of macromolecule is accounted for [134, 44]. The following Figure (8.4) explains the steps necessary for the estimation of the biosynthetic cost per molecule. Protein estimation was done with extraction of protein from lyophilysed cell using microplate assay protocol.

```
%% model, glucose -10, oxygen -1

model=readCbModel()

model.lb(findRxnIDs(model,'R_1714')) = -10; % set glucose to -10
model.lb(findRxnIDs(model,'R_1992')) = -1; % set oxygen to -1

exchangeRxns = findExcRxns(model);
media_indexes = all([exchangeRxns,model.lb ~= 0],2);
disp([model.rxns(media_indexes) model.rxnNames(media_indexes) ...
   num2cell(model.lb(media_indexes)) num2cell(model.ub(media_indexes))]);

model_solution=optimizeCbModel(model,[],'M_biomass') % f = biomass
model_solution.x(findRxnIDs(model,'R_1673')) % CO2
model_solution.x(findRxnIDs(model,'R_1862')) % ethanol
model_solution.x(findRxnIDs(model,'R_1543')) % acetat

%% model, glucose -10, oxygen -1

model=readCbModel()
model.lb(findRxnIDs(model,'R_1714')) = -10; % set glucose to -10
model.lb(findRxnIDs(model,'R_1992')) = -1; % set oxygen to -1

exchangeRxns = findExcRxns(model);
media_indexes = all([exchangeRxns,model.lb ~= 0],2);
disp([model.rxns(media_indexes) model.rxnNames(media_indexes) ...
   num2cell(model.lb(media_indexes)) num2cell(model.ub(media_indexes))]);

model_solution=optimizeCbModel(model,[],'M_biomass') %biomass flux
model_solution.x(findRxnIDs(model,'R_1673')) % CO2 production
model_solution.x(findRxnIDs(model,'R_1862')) %  ethanol
model_solution.x(findRxnIDs(model,'R_1543')) %  acetate
```

8.2.7 Biomass

Preliminary biomass estimation was based on Palsson *et al* protocol [134] and validated by experimental component analysis done with collaboration with Dr Shyam Maskapalli, University of Bath. Biorad DC protein assay was used to estimate the protein concenrations. Fatty acid profile was analysed by the University of Stirling. Total carbohydrate were estimated using Dubois *et al*, 1956 [46] protocol using 5% (w/v) aqueous phenol, set of glucose standards (0,25,50,75 and 100 µg / ml and H_2SO4 according to the established protocol.

BSA standard Promega Wizard SV Genomic DNA purification method was used to purify DNA for the extractions. Biomass estimates are corrected for the presence of moisture and minerals estimated from TGA analysis. Bligh DYER METHOD and Phenol-H2SO4 method (glucose standard) were used by Dr Shyam Maskapalli for further estimations of the biomass compositions.

The estimation of biomass is critical for simulation of flux balance analysis (FBA) under steady state where biomass equation serves as the objective function.Table **??** refers to the

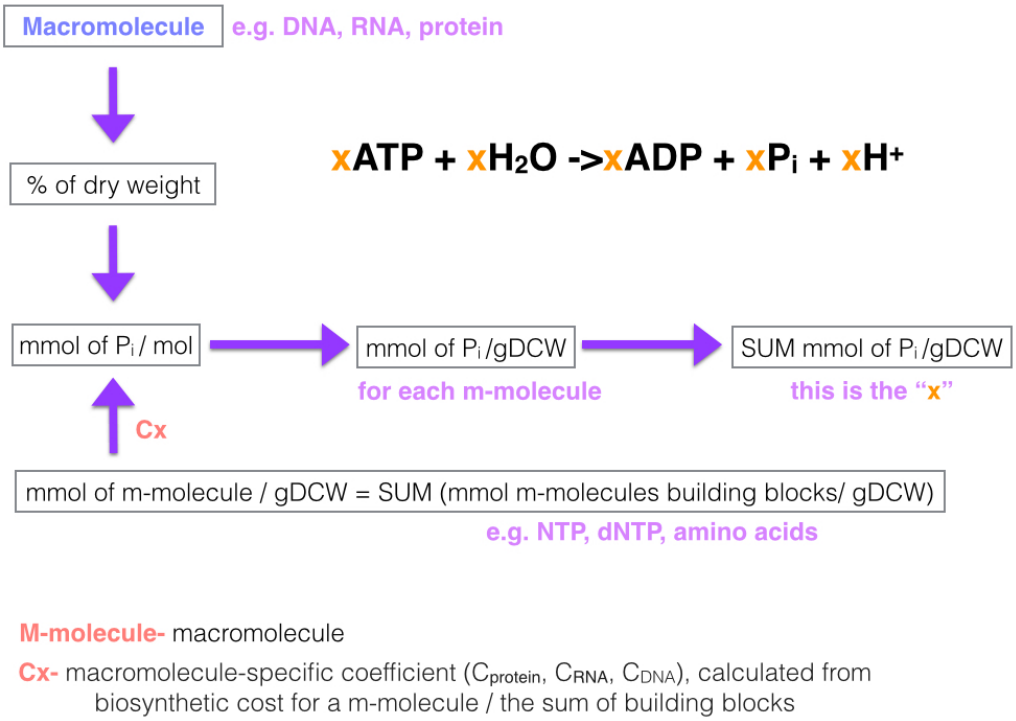


Figure 8.4: Theoretical approach for estimation of growth associated ATP maintenance [89, 134, 100].

components of biomass objective function and estimated components based on the following methodology and experimental approach.

Experimentally, the biomass composition was done in chemostat cultures by Dr Shyam Masakapalli according to Durot[47] method under the following conditions: temperature 60C, pH=6.8, aerobic and anaerobic conditions, media=ASM with 1% Glucose, total culture volume= 1.5L in 2L volume bioreactor, dilution rates: 0.05, 0.1, 0.2, 0.4 for aerobic, 0.05, 0.1 for anaerobic under low redox.

The assembled biomass reaction incorporated into the model (based on tables 4.1,4.2, 4.3, 4.4, 4.5 and 4.6) per 1 mmol per gram Dry Cell Weight, is:

0.433 glycine + 0.546 L-alanine + 0.399 L-valine + 0.369 mol L-leucine + 0.289 mol L-isoleucine + 0.192 mol L-serine + 0.285 mol L-threonine + 0.154 L-phenylalanine + 0.110 L-tyrosine + 0.109 L-tryptophan + 0.114 L-cysteine + 0.090 L-methionine + 0.336 L-lysine + 0.194 L-arginine + 0.083 L-histidine + 0.471 L-aspartate + 0.683 L-glutamate + 0.471 L-asparagine + 0.683 L-glutamine + 0.177 L-proline + 0.023 GTP + 0.023 CTP + 0.030 UTP + 0.030 dATP + 0.023 dGTP + 0.080 dCTP + 0.087 l dTTP + 0.09532 monoglu-

cosyldiacylglycerol + 0.110292 diglucosyldiacylglycerol + 0.5683 triglucosyldiacylglycerol + 0.002372 cardiolipin + 0.12084 phosphatidylglycerol + 0.04620 lysylphosphatidylglycerol + 0.3291 phosphatidylethanolmaine + 0.0079543 lipoteichoic acid (n=24) + 2.89 glycerol teichoic acid (n=45) + 2.29 glycerol teichoic acid (n=45,) + 1.47 glycerol teichoic acid (n=45) + 3.11 + 809 K + 113 Mg + 4.1 Fe(+3) + 2.9 Ca + 0.87 diphosphate + 0.289 menaquinol 7 + 0.26 10-formyltetrahydrofolate + 0.023 NAD +0.031 AMP + 0.031 ADP + 0.027 CMP + 0.009 NADP + 0.00372 CTP + 0.0007 GMP + 0.0003 GTP + 0.0004 CDP + 0.0001 NADPH + 0.0003 GDP + 108.0 mol ATP + 104.0 H2O 104.9 phosphate + 104.214 ADP + 105.0 H^+ + 1 mg biomass

The 1 mg at the end of the biomass equation denotes the production of 1 mg of cell biomass and denotes the product of all the substrate. It is in other words, production of a cell, given the quantities of its building blocks.

8.2.8 Origin of replication and terminus

The origin and terminus of replication were found using ACT software [27]. The differences in base composition in the leading and lagging strands were found by plotting GC and AT skews for the genome of *G. thermoglucosidasius* NCIMB 11955. Based on a switch in the skew (Figure 8.5), two regions on the chromosome were highlighted and assumed to correspond to the origin and terminus of replication. To confirm this we manually inspected the identified regions. Bidirectional BLAST was used to confirm the location of the OriC region, comprised of dnaN, dnaA and rmpH genes, and the gene encoding the replication termination protein 3231157 and 1336775 bases away from the arbitrarily assigned sequence start site, respectively (see Figure 8.6). Therefore, the annotated sequence has subsequently been rearranged so that 1 bp matches with the OriC region.

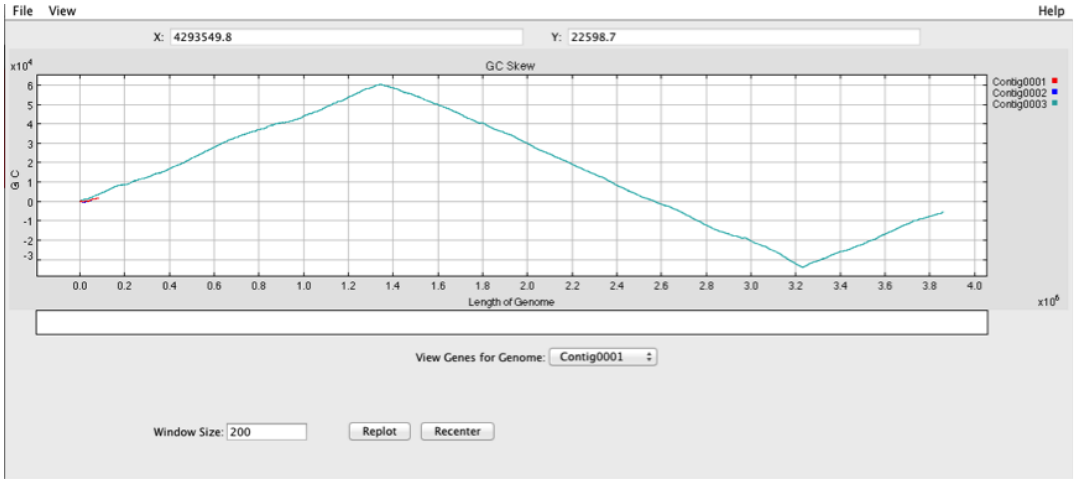


Figure 8.5: GCskew for genome of *Geobacillus thermoglucosidasius* NCIMB 11955

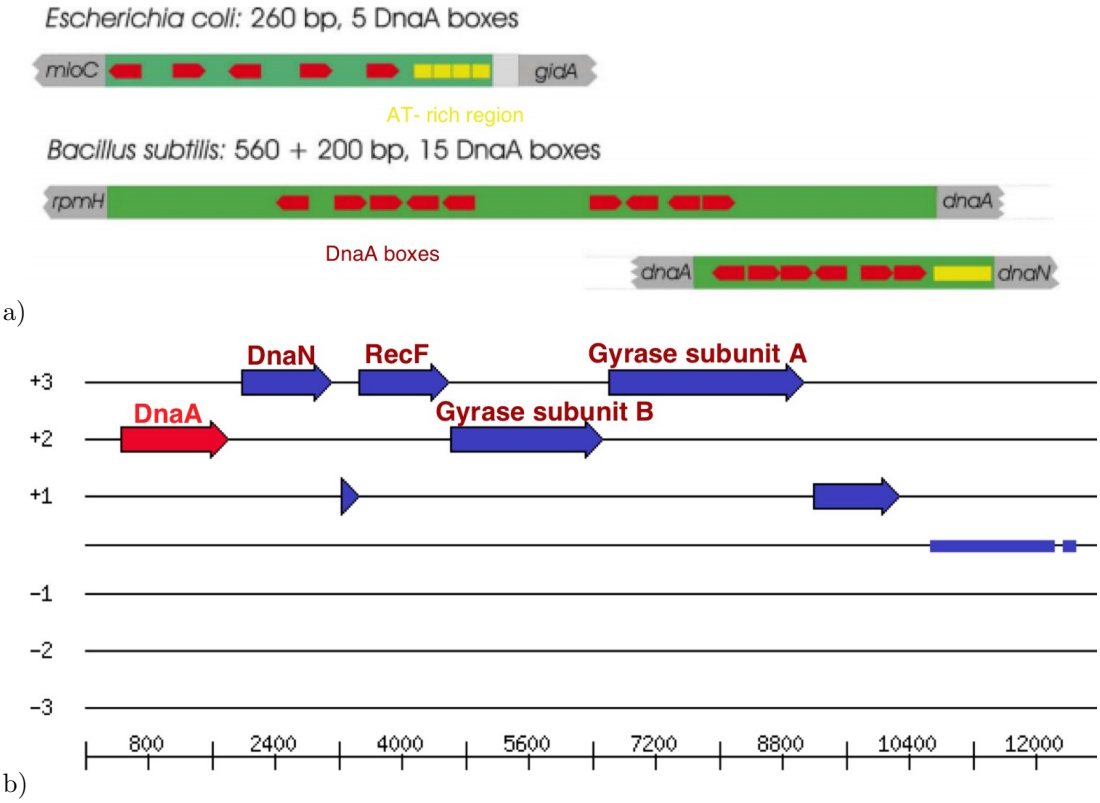


Figure 8.6: The composition of OriC region in a) *E. coli*, *Bacillus subtilis* and b) *Geobacillus thermoglucosidasius* NCIMB 11955.

8.2.9 Genome Organisation

Genome organisation analyses were done using ACT software [27]. The genome sequences were obtained from NCBIdatabases for *G. kaustophilus* HTA426 (NC_006510, GI: 56418535), *G. kaustophilus* GBlys (BASG01000001.1, GI: 514407193) *G. thermodenitrificans* NG8O-2 (NC_009328.1, GI:138893679), *G. thermoleovorans* CCB-US.3-UF5 (NC_016593.1, GI: 375006802), *G. thermoglucosidasius* C56-YS93 (NC_015660.1 GI:336233546), *G. stearothermophilus* ATCC 7953 (NZ_JALS01000119.1, GI:696478034),*G. thermoglucosidans* TNO-09.020 (CM001483.1, GI: 384525213) *Bacillus megaterium* ATCC 14581 (CP001982.1, GI:294799901), *Bacillus subtilis* 168 (NC_000964.3, GI: 255767013). *MEGA6* software [129] was used to produce phylogenetic tree for genus *Geobacillus* using *recN* gene sequences obtained from UniProt and NCBI website, with the following accession numbers can be found in the table 8.5.

Table 8.5: Bacterial strains and *recN* gene accession number used for the generation of phylogenetic tree.

Strain	GenBank/EMBL/DDBJ accession number
G. caldotenax BGSC 96A4	AY609005
Geobacillus sp. CAMR5420	A0A063YTE8
G. vulcani BGSC 97A1	AY609007
*G. litanicus*BGSC W9A89	Q6E1S8
Geobacillus sp. NTU 03	F1K0Y8
G. kaustophilus BGSC 90A1	AY609001
G. thermoleovorans BGSC 96A1	AY609003
G. kaustophilus HTA426	Q5L3Y9
G. thermoleovorans CCB US3 UF5	AY609006
Geobacillus sp. A8	T0NXL7
Geobacillus sp. MAS1	V6VD46
Geobacillus sp. WSUCF1	S7UBR5
Geobacillus sp. C56-T3	ADI26142
Geobacillus sp. Y412MC52	ADU94794
Geobacillus sp. Y412MC61	ACX76952
G. thermocatenulatus BGSC 93A1	AY609002
Geobacillus sp. G1w1	KFX34333
G. stearothermophilus BGSC 9A20	AY608997

Table 8.5: Bacterial strains and *recN* gene accession number used for the generation of phylogenetic tree.

Strain	GenBank/EMBL/DDBJ accession number
G. stearothermophilus BGSC W9A 29	AY608998
G. stearothermophilus BGSC 9A2	AY609030
G. subterraneus BGSC 91A2	AY609023
G. uzensis BGSC 91A2	ACX76952
G. kaue BGSC W9A78	AY609042
Geobacillus sp. JF8	AGT32726
G. thermodenitrificans BGSC 94A1	AY608960
G. thermodenitrificans BGSC W9A26	AY608963
G. caldoxylosilyticus BGSC W98A1	AY609016
G. caldoxylosilyticus BGSC W9A36	AY609017
G. "stearothermophilus" NUB 3621	AY609015
G. toebii BGSC W9A22	AY609049
Geobacillus sp. WCH70	ACS25020
G. toebii BGSC 99A1	AY608982
Bacillus "thermoantarcticus" BGSC 20A1	AY609047
G. thermoglucosidasius NBRC 107763	GAJ42512
G. thermoglucosidasius BGSC 95A1	AY608981
G. thermoglucosidasius NCIMB 11955	this study
G. thermoglucosidans TNO-09.020	EID45490
Geobacillus sp. Y4.1MC1	ADP73991
G. thermoglucosidasius C56-YS93	A0A0M1QJZ6
Bacillus subtilis str 168	QU35_13255

The evolutionary history was inferred using the Neighbor-Joining method [108]. The evolutionary distances were computed using the Maximum Composite Likelihood method [5] and are in the units of the number of base substitutions per site. The analysis involved 13 nucleotide sequences. Codon positions included were 1st+2nd+3rd+stop. All positions containing gaps and missing data were eliminated. There were a total of 1718 positions in the final dataset. Evolutionary analyses were conducted in *MEGA6* [129].

8.2.10 PHAST

The *PHAST* [149] software was used to determine if the type-strains of each clade has intact or incomplete phage sequence present. The bacterial strains analysed were: *G. kaustophilus* HTA426 (NC_006510, GI: 56418535), *G. kaustophilus* GBlys (BASG01000001.1, GI: 514407193) *G. thermodenitrificans* NG8O-2 (NC_009328.1, GI:138893679), *G. thermoleovorans* CCB-US.3-UF5 (NC_016593.1, GI: 375006802), *G. thermoglucosidasius* C56-YS93 (NC_015660.1, GI:336233546), *G. stearothermophilus* ATCC 7953 (NZ_JALS01000119.1, GI:696478034). The software was used with accordance to the its manual.

8.2.11 RNA-SeqAnalysis

The transcriptome data was generated by Dr Leann Bacon and Dr Shyam Masakapalli at the University of Bath in the group of Professor David Leak. The cells were grown under anaerobic and aerobic conditions, in rich and minimal media. The data was done in duplicates. The cells were grown in the following conditions:

- rich media under aerobic growth
- rich media under anaerobic growth
- minimal media under aerobic growth
- minimal media under anaerobic growth

The RNeasy mini kit (Qiagen) was used to extract the least degraded RNA. The samples were then send for analysis to Deep Seq: Next Generation Sequencing Facility at the University of Nottingham. The following analysis of the reads was done by the author. CLC genomics was used to align the transcripts to the annotated genome of *G. thermoglucosidasius* NCIMB 11955. *DESeq2* [3] was used for the RNA-Seq data analysis according to the protocol [77], combined with *GAGE* tool [79] workflow in R/Bioconductor environment. *DESeq2* was used for the differential expression analysis. The *GAGE* and *Pathview* [78] were used for the pathway analysis and visualisation. Both up-regulated and down-regulated genes were visualised. The script generated by the author for this analysis is attached.

```
##### FOR RNAseqData
library("DESeq2")
 table<-read.csv("RNAaseqR1an2R2An2R2A4R1A42.csv", header=TRUE, row.names=1)
 head(table)
      R1AN2 R2An2 R2A4 R1A4
peg0001  1265  2202 1025 1736
peg0003   206   461 2103  397
peg0004   464   500  589 1663
peg0005   160   424  549  905
peg0006   104   291 1612  497
peg0008   516  1238 2667  886


samples <- data.frame(row.names=c("R1AN2","R2An2","R2A4","R1A4"),
condition=as.factor(c(rep("Anaerobic",2),rep("Aerobic",2))))
samples
    condition
R1AN2 Anaerobic
R2An2 Anaerobic
R2A4    Aerobic
R1A4    Aerobic


CDS<-DESeqDataSetFromMatrix(countData=table, colData=samples, design=~condition)
CDS_1<-DESeq(CDS)
estimating size factors
estimating dispersions
gene-wise dispersion estimates
mean-dispersion relationship
final dispersion estimates
fitting model and testing
res<-results(CDS_1)
 head(res)
log2 fold change (MAP): condition Anaerobic vs Aerobic
Wald test p-value: condition Anaerobic vs Aerobic
DataFrame with 6 rows and 6 columns
       baseMean log2FoldChange     lfcSE      stat     pvalue      padj
      <numeric>      <numeric> <numeric>  <numeric>  <numeric> <numeric>
peg0001 1595.3071      1.0959563 0.5847466  1.8742415 0.06089714 0.2096244
peg0003  703.6078     -1.1729472 0.7881943 -1.4881448 0.13671271 0.3409587
peg0004  729.1136     -0.1859726 0.6810594 -0.2730637 0.78480427 0.8956546
peg0005  430.1017     -0.5721169 0.5950516 -0.9614577 0.33632210 0.5653072
peg0006  529.9389     -1.6911848 0.7323551 -2.3092417 0.02093017 0.1030333
peg0008 1220.3069     -0.3834418 0.7095279 -0.5404182 0.58890866 0.7707517

>  write.csv(res,file="RNAseq001results.csv")

 resOrdered<-res[order(res$padj),]
 summary(res)
    Length       Class        Mode
         6 DESeqResults          S4

 plotMA(res, ylim=c(-7,7),main="DESeq2")

> select<-order(rowMeans(counts(CDS_1, normalized=TRUE)), decreasing=TRUE)[1:30]
 hmcol<-colorRampPalette(brewer.pal(9,"GnBu"))(100)
 heatmap.2(counts(CDS_1, normalized=TRUE)[select,], col=hmcol,Rowv=FALSE, Colv=FALSE,
scale="none", dendrogram="none", trace="none", margin=c(10,6))
```

```
 resSig<-subset(resOrdered, padj<0.1)

> write.csv(as.data.frame(resSig), file="Condition_treated_resulst_01.csv")
 rld<-rlog(CDS_1)
 vsd<-varianceStabilizingTransformation(CDS_1)

>par(mfrow=c(1,3))
> notAllZero<-(rowSums(counts(CDS_1))>0)
> meanSdPlot(log2(counts(CDS_1, normalized=TRUE)[notAllZero,]+1))
> meanSdPlot(assay(rld[notAllZero,]))
> meanSdPlot(assay(vsd[notAllZero,]))


##### HeatMap

heatmap.2(assay(rld)[select,], col=hmcol, Rowv=FALSE, Colv=FALSE, scale="none",
dendrogram="none", trace="none", margin=c(5,30))
heatmap.2(counts(CDS_1, normalized=TRUE)[select,], col=hmcol, Rowv=FALSE, Colv=FALSE,
scale="none", dendrogram="none", trace="none", margin=c(5,30))
heatmap.2(assay(vsd)[select,], col=hmcol, Rowv=FALSE, scale="none", dendrogram="none",
trace="none", margin=c(5,30))

###Histogram of p values for all tests
> use<-res$baseMean > attr(res,"filterThreshold")
> table(use)
use
FALSE  TRUE
  186  3518
> resFilt<-res[use & !is.na(res$pvalue),]
> orderInPlot<-order(resFilt$pvalue)
> showInPlot<- (resFilt$pvalue[orderInPlot] <= 0.08)
> alpha<-0.1

#### DESeq2 visualisation using GAGE Pathway view

deseq2.res<-results(CDS_1)

#direction of fc, depends o the levels, the first level taken as the reference and the second as
experiment

deseq2.fc=deseq2.res$log2FoldChange

names(deseq2.fc)=rownames(deseq2.res)
exp.fc=deseq2.fc

out.suffix="deseq2"

##GAGE for pathway analysis. I used G.thermoglucosidasius C56-Y.. as KEGG ids

require(gage)
data(kegg.gs)

##get gene set specific data for G. thermoglucosidasius, gene type= "kegg id"

kg.gth=kegg.gsets("gth", id.type="kegg")
names(kg.gth)
lapply(kg.gth, head, 3)
```

```
$kg.sets
$kg.sets$`gth00010 Glycolysis / Gluconeogenesis`
 [1] "Geoth_0237" "Geoth_0238" "Geoth_0239" "Geoth_0241" "Geoth_0268" "Geoth_0442"
"Geoth_0443"
 [8] "Geoth_0444" "Geoth_0445" "Geoth_0446" "Geoth_0632" "Geoth_0652" "Geoth_0811"
"Geoth_0855"
[15] "Geoth_0866" "Geoth_0879" "Geoth_0897" "Geoth_0898" "Geoth_0910" "Geoth_1233"
"Geoth_1307"
[22] "Geoth_1572" "Geoth_1595" "Geoth_1596" "Geoth_1597" "Geoth_1598" "Geoth_1628"
"Geoth_1917"
[29] "Geoth_2081" "Geoth_2227" "Geoth_2267" "Geoth_2349" "Geoth_2366" "Geoth_2367"
"Geoth_2368"
[36] "Geoth_2478" "Geoth_2479" "Geoth_2480" "Geoth_2860" "Geoth_2861" "Geoth_2862"
"Geoth_2863"
[43] "Geoth_2925" "Geoth_2974" "Geoth_3108" "Geoth_3118" "Geoth_3288" "Geoth_3351"
"Geoth_3494"
[50] "Geoth_3520" "Geoth_3524" "Geoth_3823" "Geoth_3831" "Geoth_3834" "Geoth_3880"
"Geoth_3897"

$kg.sets$`gth00020 Citrate cycle (TCA cycle)`
 [1] "Geoth_0237" "Geoth_0238" "Geoth_0239" "Geoth_0811" "Geoth_0902" "Geoth_0903"
"Geoth_0904"
 [8] "Geoth_0968" "Geoth_0969" "Geoth_0970" "Geoth_1307" "Geoth_1595" "Geoth_1596"
"Geoth_1597"
[15] "Geoth_1598" "Geoth_2366" "Geoth_2367" "Geoth_2368" "Geoth_2478" "Geoth_2479"
"Geoth_2480"
[22] "Geoth_2552" "Geoth_2618" "Geoth_2619" "Geoth_2713" "Geoth_2714" "Geoth_2837"
"Geoth_2860"
[29] "Geoth_2861" "Geoth_2862" "Geoth_2863" "Geoth_2896" "Geoth_2897" "Geoth_3395"
"Geoth_3444"
[36] "Geoth_3509"

$kg.sets$`gth00030 Pentose phosphate pathway`
 [1] "Geoth_0067" "Geoth_0234" "Geoth_0632" "Geoth_0897" "Geoth_1177" "Geoth_1333"
"Geoth_1335"
 [8] "Geoth_1353" "Geoth_2081" "Geoth_2256" "Geoth_2257" "Geoth_2314" "Geoth_2567"
"Geoth_2742"
[15] "Geoth_3109" "Geoth_3110" "Geoth_3116" "Geoth_3117" "Geoth_3118" "Geoth_3288"
"Geoth_3815"
[22] "Geoth_3831" "Geoth_3833" "Geoth_3834"


$sigmet.idx
[1] 78 81 82

$sig.idx
[1] 78 81 82

$met.idx
[1] 1 2 3

$dise.idx
integer(0)


#save the gene set data for future use
save(kg.gth, file="kg.gth.RData")

#I want to use metabolic pathways in your analysis. So extract those pathways for my analysis:
```

```
kegg.gs=kg.gth$kg.sets[kg.gth$sigmet.idx]

#Call GAGE with:
fc.kegg.p <- gage(exp.fc, gsets = kegg.gs, ref = NULL, samp = NULL)

deseq2.kegg.sig<-sigGeneSet(fc.kegg.p, outname="sig.kegg",pdf.size=c(7,8))

[1] "No heatmap produced for up- or down-regulated gene sets, only 1 or none signficant."
[1] "there are 0 signficantly up-regulated gene sets"
[1] "there are 1 signficantly down-regulated gene sets"

####code for up-regulated pathways
sel <- fc.kegg.p$greater[, "q.val"] < 0.1 & !is.na(fc.kegg.p$greater[, "q.val"])
path.ids <- rownames(fc.kegg.p$greater)[sel]
sel.l <- fc.kegg.p$less[, "q.val"] < 0.1 &!is.na(fc.kegg.p$less[,"q.val"])
path.ids.l <- rownames(fc.kegg.p$less)[sel.l]
path.ids2 <- substr(c(path.ids, path.ids.l), 1, 8)
require(pathview)
#view first 3 pathways as demo
pv.out.list <- sapply(path.ids2[1:3], function(pid) pathview(gene.data =  exp.fc, pathway.id = pid,
species = "gth", out.suffix=out.suffix, gene.idtype="KEGG"))
```

Bibliography

[1] Larhlimi A. and A. Bockmayr. A new constraint-based description of the ready-state flux cone of the metabolic networks. *Discrete Applied Mathematics*, 157:2257–2266, 2009.

[2] A A Aghapour, G Moussavi, and K Yaghmaeian. Biological degradation of catechol in wastewater using the sequencing continuous-inflow reactor (scr). *J. Envron Health Sci Eng*, 11(3), 2013.

[3] S Anders and W Huber. Differential expression analysis for sequence count data. *Genome Biol*, 11(10):1465–9606, 2010.

[4] L L Anzaldi, , and E P Skaar. Overcoming the heme paradox: Heme toxicity and tolerance in bacterial pathogens. *Infec and Immun*, 78(12):4977–89, 2010.

[5] M L Artiguez and I M de Maranon. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc of Nat Acad of Sciences (USA)*, 101:11030–11035, 2004.

[6] M L Artiguez and I M de Maranon. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc of Nat Acad of Sciences (USA)*, 101:11030–11035, 2004.

[7] M L Artiguez and I M de Maranon. Inactivation of spores and vegetative cells of *Bacillus subtilis* and *Geobacillus stearothermophilus* by pulsed light. *Innov Food Sc and Em Tech*, 28:52–58, 2015.

[8] G Ashwell, A J Wahba, and J Hickman. Uronic acid metabolism in bacteria. i. purification and properties of uronic acid isomerase in *Escherichia coli*. *J. Biol. Chem.*, 235(6):1559–1565, 1960.

[9] T Auer, J J Sninsky, D H Gelfand, and T W Myers. Selective amplification of rna utilizing the nucleotide analog ditp and thermus thermophilus dna polymerase. *Nucleic Acids Res*, 24:5021–5025, 1996.

[10] K Awwad, A Desai, C Smith, and M Sommerhalter. Structural and functional characterization of a noncanonical nucleoside triphosphate pyrophosphatase from *Thermotoga maritima*. *Acta Cryst*, D69:184–193, 2013.

[11] R Aziz, D Bartels, A A Best, M DeJongh, T Disz, R A Edwards, K Formsma, S Gerdes, E M Glass, M Kubal, M Meyer, G J Olsen, R Olson, A L Osterman, R A Overbeek, L K McNeil, D Paarmann, T Paczian, B Parrello, G D Pusch, C Reich, R Stevens, O Vassieva, V Vonstein, A Wilke, and O Zagnitko. The rast server: Rapid annotations using subsystems technology. *BMC Genomics*, 9(75), 2008.

[12] J. E. Bailey. Mathematical modelling and analysis in biochemical engineering: past accomplishments and future opportunities. *Biotechnology Progress*, 14:8–20, 1998.

[13] S Becker and O Palsson, B. Context-specific metabolic networks are consistent with experiments. *PLoS Comput Biol*, 4(5), 2008.

[14] National Center for Biotechnology Information Bethesda (MD): National Library of Medicine (US). The ncbi handbook [internet]. 2002.

[15] T A Bobik, G D Havemann, R J Busch, D S Williams, and H C Aldrich. The propanediol utilization (pdu) operon of *Salmonella enterica serovar Typhimurium* lt2 includes genes necessary for formation of polyhedral organelles involved in coenzyme b(12)-dependent 1, 2-propanediol degradation. *J Bacteriol*, 181(19):5967–75, 1999.

[16] T A Bobik, Xu Y, R M Jeter, K E Otto, and J R Roth. Propanediol utilization genes (pdu) of *Salmonella typhimurium*: Three genes for the propanediol dehydratase. *J. Bacteriol.*, 179(21), 1997.

[17] B K Bonde, D J V Beste, E Laing, M Kierzek, A, and J McFadden. Differential producibility analysis (dpa) of transcriptomic data with metabolic networks: Deconstructing the metabolic response of *M. tuberculosis*. *PLoS Comput Biol*, 7(6), 2011.

[18] A. Bookstein, V. Kulyukin, and T. Raita. Generalized hamming distance. *Information Retrieval*, 5(4):353–375, 2002.

[19] R T Bossi, A Negri, G Tedeschi, and A Mattevi. Structure of fad-bound l-aspartate oxidase: insight into substrate specificity and catalysis. *Biochemistry*, 41(9):3018–24, 2002.

[20] H Brussow, C Canchaya, and W D. Hardt. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev*, 68(3):560–602, 2004.

[21] B B Buchanan and M C Evans. The synthesis of alpha-ketogutarate from succinate and carbon dioxide by a subcellular preparation of a photosynthetic bacterium. *Proc. Natl. Acad. Sci. U.S.A*, 54(4):1212–8, 1965.

[22] M Bulmer. The selection-mutation-drift theory of synonymous codon usage. *Genetics*, 129(3):897–907, 1991.

[23] Wagner C. and R. Urbanczyk. The geometry of the flux cone of a metabolic network. *Biophysical Journal*, 89:3837–3845, 2005.

[24] B J Campbell and S C Cary. Abundance of reverse tricarboxylic acid cycle genes in free-living microorganisms at deep-sea hydrothermal vents. *App Env Micro*, 70(10), 2004.

[25] C Canchaya, G Fournous, S Chibani-Chennoufi, M L Dillmann, and Brussow H. Phage as agents of lateral gene transfer. *Curr Opin Microbiol*, 6(4):417–24, 2003.

[26] Y Cao, R Zhang, C Sun, E Cheng, Y Liu, and M Xian. Fermentative succinate production: An emerging technology to replace the traditional petrochemical processes. *Biomed Res.*, 2013, 2013.

[27] T J Carver, K M Rutherford, M Berriman, M A Rajandream, B G Barrell, and Parkhill. *act*: the artemis comparison tool. *Bioinformatics*, 21(16):3422–3, 2005.

[28] T H Chiu and D S Feingold. L-rhamnulose 1-phosphate aldolase from escherichia coli. crystallization and properties. *Biochemistry*, 8(1):98–108, 1969.

[29] H Chu, B Xin, P Liu, Y Wang, L Li, X Liu, X Zhang, C Ma, P Xu, and C Gao. Metabolic engineering of escherichia coli for production of (2s,3s)-butane-2,3-diol from glucose. *Biotech for Biofuels*, 8(143), 2006.

[30] J H Chung, J H Back, Y I Park, and Y S Han. Biochemical characterization of a novel hypoxanthine/xanthine dntp pyrophosphatase from *Methanococcus jannaschii*. *Nucleic Acids Res*, 29(14):3099–107, 2001.

[31] A C Cihan, B Ozcan, N Tekin, and C Cokmus. *Geobacillus thermodenitrificans* subsp. calidus, subsp. nov., a thermophilic and α-glucosidase producing bacterium isolated from kizilcahamam, turkey. *Extremophiles*, 6(2), 2002.

[32] C Claudel-Renard, C Chevalet, T Faraut, and D Kahn. Enzyme-specific profiles for genome annotation: Priam. *Nucleic Acids Res.*, 31:66336639, 2003.

[33] S Clermont, C Corbier, Y Mely, D Gerard, A Wonacott, and G Branlant. Determinants of coenzyme specificity in glyceraldehyde-3-phosphate dehydrogenase: role of the acidic residue in the fingerprint region of the nucleotide binding fold. *Biochemistry*, 32(38):10178–84, 1993.

[34] G N Cohen. *Microbial biochemistry.* Springer Netherlands, 2014.

[35] C Colijn, A Brandes, J Zucker, DS Lun, B Weiner, M R Farhat, T Y Cheng, D B Moody, M Murray, and J E Galagan. Interpreting expression data with metabolic flux models: predicting mycobacterium tuberculosis mycolic acid production. *PLoS Comput Biol.*, 5(8), 2009.

[36] C Corbier, S Clermont, P Billard, T Skarzynski, C Branlant, A Wonacott, and G. Branlant. Probing the coenzyme specificity of glyceraldehyde-3-phosphate dehydrogenases by site-directed mutagenesis. *Biochemistry*, 29(30):7101–6, 1990.

[37] D N Correa-Llanten, S A Munoz-Ibacache, M E Castro, P A Munoz, and J B Blamey. Gold nanoparticles synthesized by *Geobacillus* sp. strain id17 a thermophilic bacterium isolated from deception island, antarctica. *Microb Cell Fact.*, 12(75), 2013.

[38] R E Cripps, K Eley, D J Leak, B Rudd, M Taylor, M Todd, S Boakes, S Martin, and T Atkinson. Metabolic engineering of geobacillus thermoglucosidasius for high yield ethanol production. *Metabolic Engineering*, 11(6):398–408, 2009.

[39] C S Crowley, M R Sawaya, T A Bobik, and Yeates T O. Structure of the pduu shell protein from the pdu microcompartment of *Salmonella*. *Structure*, 16(9):1324–32, 2008.

[40] M A Cynkin and G Ashwell. Uronic acid metabolism in bacteria. *J. Biol. Chem.*, 235(6):1576–1579, 1959.

[41] Segre D, Zucker J, Katz J, Lin X, D'haeseleer P, Rindone WP, Kharchenko P, Nguyen DH, Wright MA, and Church GM. From annotated genomes to metabolic flux models and kinetic parameter fitting. *OMICS; A journal of Integrative Biology*, 7(3), 2003.

[42] Segr D. From annotated genomes to metabolic flux models and kinetic parameter fitting. *OMICS; A journal of Integrative Biology*, 7(3), 2003.

[43] T A Dailey, T O Boynton, A N Albetel, S Gerdes, M K Johnson, and H A Dailey. Discovery and characterization of hemq: an essential heme biosynthetic pathway component. *J Biol Chem*, 285(34):25978–86, 2010.

[44] D M Dauer, M W Luckenbach, and A J Rodi Jr. Abundance biomass comparison (abc method): effects of an estuarine gradient, anoxic/hypoxic events and contaminated sediments. 270(7):1474–82, 1993.

[45] P De Maayer, P J Brumm, Mead D A, and Cowan D A. Comparative analysis of the geobacillus hemicellulose utilization locus reveals a highly variable target for improved hemicellulolysis. *BMC Genomics*, 15(836), 2014.

[46] M Dubois. Colorimetric method for determination of sugars and related substances. *Analytical Chemistry*, 28(3), 1956.

[47] M Durot, P Y Bourguignon, and V Schachter. Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiol Rev*, 33(1):164–90, 2009.

[48] J.S. Edwards and B.O Palsson. The /textitEscherichia coli mg1655 in silico metabolic genotype: Its definition, characteristics, and capabilities. *PNAS*, 97(10), 2000.

[49] V Egelhofer, I Schomburg, and D Schomburg. Automatic assignment of ec numbers. *PloS Comput Biol*, 6(1), 2010.

[50] A M Feist and B O Palsson. The biomass objective function. *Curr Opin Microbiol*, 13(3):344–349, 2010.

[51] L Feng, C Stathopoulos, I Ahel, A Mitra, D Tumbula-Hansen, T Hartsch, and D Soll. Aminoacyl-trna formation in the extreme thermophile *Thermus thermophilus*. *Extremophiles*, 6(2), 2002.

[52] L Feng, W Wang, J Cheng, Y Ren, G Zhao, C Gao, Y Tang, X Liu, W Han, X Peng, R Liu, and L. Wang. Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* ng80-2 isolated from a deep-subsurface oil reservoir. *Proc Natl Acad Sci USA*, 104(13):5602–7, 2007.

[53] S Fillinger, S Boschi-Muller, S Azza, E Dervyn, G Branlant, and S Aymerich. Two glyceraldehyde-3-phosphate dehydrogenases with opposite physiological roles in a nonphotosynthetic bacterium. *J Biol Chem*, 275(19):14031–7, 2000.

[54] A Flamholz, E Noor, A Bar-Even, Y Lubling, D Davidi, and R Milo. An integrated open framework for thermodynamics of reactions that combines accuracy and coverage. *Bioinformatics*, 28:2037–44, 2012.

[55] A Flamholz, E Noor, A Bar-Even, and R Milo. *equilibrator* - the biochemical thermodynamics calculator. *Nucleic Acid Res*, 40:D770–5, 2012.

[56] R W Hamming. Error detectin and error correcting codes. *The Bell Sys Tech Jour*, 24(2), 1950.

[57] M Hansson and L Hederstedt. Cloning and characterization of the *Bacillus subtilis* hemehy gene cluster, which encodes protoheme ix biosynthetic enzymes. *J Bacteriol*, 174(24):8081–93, 1992.

[58] C S Henry, J F Zinner, M P Cohoon, and R L Stevens. ibsu1103: a new genome-scale metabolic model of bacillus subtilis based on seed annotations. *Genome Biol*, 10(69), 2009.

[59] S L Hii, J S Tan, T C Ling, and A B Ariff. Pullulanase: Role in starch hydrolysis and potential industrial applications. *Enzyme Research*, 2012:1–14, 2012.

[60] A H Hussein, B K Lisowska, and Leak D J. The genus geobacillus and their biotechnological potential. *Adv Appl Microbiol.*, 92:1–48, 2015.

[61] Schellenberger J, R Que, R M Fleming, I Thiele, J D Orth, A M Feist, D C Zielinski, A Bordbar, N E Lewis, S Rahmanian, J Kang, D R Hyduke, and B O Palsson. Quantitative prediction of cellular metabolism with constraint-based models: the cobra toolbox v2.0. *Nature Prot*, 6:1290–1307, 2011.

[62] XJ Ji, H Huang, and Ouyang PK. Microbial 2,3-butanediol production: a state-of-the-art review. *Biotechnol Adv.*, 29:351364, 2015.

[63] J J Just, C E M Stevenson, L Bowater, A Tanner, D M Lawson, and Bornemann S. A closed conformation of *Bacillus subtilis* oxalate decarboxylase oxdc provides evidence for the true identity of the active site. *J Biol Chem*, 279:19867–74, 2004.

[64] M et al. Kanehisa. Kegg for integration and interpretation of large-scale molecular datasets. *Nucleic Acids Research*, 40:D109–D114, 2012.

[65] Goto S. Kanehisa M. Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28:27–30, 2000.

[66] T Kitatani, Y Nakamura, K Wada, T Kinoshita, M Tamoi, S Shigeoka, and T Tadab. Structure of nadp-dependent glyceraldehyde-3-phosphate dehydrogenase from *Synechococcus* pcc7942 complexed with nadp. *Acta Crystallogr Sect F Struct Biol Cryst Commun.*, 62(4):315–319, 2006.

[67] E Klipp, W Liebermeister, C Wierling, A Kowald, H Lehrach, and R Herwig. *Systems Biology.* Wiley-Blackwell, Verlag GmbH & KGaA, Weinheim, 1993.

[68] M1 Lakshmanan, K Yu, L Koduru, and D Y Lee. In silico model-driven cofactor engineering strategies for improving the overall nadp(h) turnover in microbial cell factories. *J Ind Microbiol Biotechnol.*, 10:1401–14, 2015.

[69] M Latendresse. Efficiently gap-filling reaction networks. *BMC Bioinformatics*, 15(225), 2014.

[70] M Latendresse, M Krummenacker, M Trupp, and P D Karp. Construction and completion of flux balance models from pathway databases. *Bioinformatics*, 28(3), 2012.

[71] H Lehnherr, E Maguin, S Jafri, and M B. Yarmolinsky. Plasmid addiction genes of bacteriophage p1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *J Mol Biol.*, 233(3):414–28, 1993.

[72] N E Lewis, H Nagarajan, and B O Palsson. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silica methods. *Nature Reviews*, 10:291–305, 2012.

[73] R Liberal and J W Pinney. Simple topological properties predict functional misannotations in a metabolic network. *Bioinformatics*, 29(13), 2013.

[74] R F Liberal, B K Lisowska, J Pinney, and D J. Leak. Pathwaybooster: a tool to support the comparison and curation of metabolic models.(plus supplementary information). *BMC Bioinformatics*, 16(86), 2015.

[75] L. Liu, R. Agren, S. Bordel, and J. Nielsen. Use of genome-scale metabolic models for understanding microbial physiology. *FEBS Letters*, 584:2556–2564, 2010.

[76] C Lopez-Aguilar, J A Ruiz-Maso, T S Rubio-Lepe, M Sanz, and G del Solar. Translation initiation of the replication initiator repb gene of promiscuous plasmid pmv158 is led by an extended non-sd sequence. *Plasmid*, 70(1):69–77, 2013.

[77] W Luo. $rna-seq$ and $gene-set\ analysis$ workflows. *tool documentation*, 2014.

[78] W Luo and C Brouwer. *pathview*: and *r/bioconductor* package for pathway-based data integration and visualisation. *Bioinformatics*, 29(14):1830–1831, 2013.

[79] W Luo, M Friedman, K Shedden, K Hankenson, and P Woolf. *gage*: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*, 10(1), 2009.

[80] Peregrin-Alvarez J M, C Sanford, and J Parkinson. The conservation and evolutionary modularity of metabolism. *Genome Biology*, 10(R63), 2009.

[81] D Maglott, J Ostell, K D Pruitt, and T Tatusova. *entrezgene*: gene-centered information at ncbi. *Nucleic Acids Res.*, 33:D54–D58, 2005.

[82] R Marchant and I M Banat. The genus *Geobacillus* and hydrocarbon utilization. *Handbook of Hydrocarbon and Lipid Microbiology*, 2010.

[83] I Marinoni, S Nonnis, C Monteferrante, P Heathcote, E Hartig, L H Bottger, A X Trautwein, A Negri, A M Albertini, and G Tedeschi. Characterization of l-aspartate oxidase and quinolinate synthase from *Bacillus subtilis*. *FEBS J.*, 275(20):5090–107, 2008.

[84] W A Meloche, H Pand Wood. The mechanism of 2-keto-3-deoxy-6-phosphogluconate aldolase. 3. nature of the inactivation by fluorodinitrobenzene. *J. Biol. Chem.*, 246:4028–4035, 1971.

[85] P Mendes, N J Stanford, and K Smallbone. Kinetic modelling of large-scale metabolic networks. *Proceedings of the 9th International Conference on Computational Methods in Systems Biology*, pages 5–6, 2011.

[86] E Miclet, V Stoven, P A Michels, F R Opperdoes, J Y Lallemand, and F Duffieux. Nmr spectroscopic analysis of the first two steps of the pentose-phosphate pathway elucidates the role of 6-phosphogluconolactonase. *J Biol Chem*, 276(37):34840–6, 2001.

[87] G Molla, L Motteran, V Job, M S Pilone, and Pollegioni L. Kinetic mechanisms of glycine oxidase from bacillus subtilis. *Arch Microbiol*, 133:300–302, 2003.

[88] N Nakayama, I Narumi, S Nakamoto, and H Kihara. Complete nucleotide sequence of pstk1, a cryptic plasmid from *Bacillus stearothermophilus* tk015. *Biotech. Let.*, 15(10):1013–6, 1993.

[89] F C Neidhardt. Escherichia coli and salmonella: cellular and molecular biology. *ASM Press*, 1996.

[90] H Niu, D Leak, N Shah, and C Kontoravdi. Metabolic characterization and modeling of fermentation process of an engineered *Geobacillus thermoglucosidasius* strain for bioethanol production with gas stripping. *Chem Eng Sc*, 122(27):138–149, 2015.

[91] E Noor, H S Haraldsdottir, R Milo, and R M T Fleming. Consistent estimation of gibbs energy using component contributions. *PLoS Comput Biol*, 9, 2013.

[92] E J O'Brien, J A Lerman, R L Chang, D R Hyduke, and B O Palsson. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol Syst Biol*, 9(693), 2013.

[93] A Osterman and R Overbeek. Missing genes in metabolic pathways: a comparative genomics approach. *Current Opinion in Chemical Biology*, 7:238–251, 2003.

[94] S Pabinger, R Rader, R Agren, J Nielsen, and Z Trajanoski. *memosys*: Bioinformatics platform for genome-scale metabolic models. *BMC Systems Biology*, 5(1), 2011.

[95] S Pabinger, R Snajder, T Hardiman, M Willi, A Dander, and Z Trajanoski. *memosys* 2.0: an update of the bioinformatics database for genome-scale models and genomic data. *Database (Oxford)*, 2014, 2014.

[96] Bernhard O. Palsson. *Systems Biology: Simulation of Dynamic Network States*. Cambridge University Press, 2011.

[97] E Pedone, D Limauro, and S Bartolucci. The machinery for oxidative protein folding in thermophiles. *Antioxid Redox Signal*, 10(1):157–69, 2008.

[98] P P Peralta-Yahya and Keasling J D. Advanced biofuel production in microbes. *Biotechnol. J.*, 5:147–162, 2010.

[99] J Perez, J Munoz-Dorado, T de la Rubia, and J Martnez. Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. *Int Microbiol.*, 5(2):52–63, 2002.

[100] S J Pirt. Maintenance energy: a general model for energy-limited and energy-sufficient growth. *Eur J Biochem*, 270(7):1474–82, 1982.

[101] G Piskachev, W Ali, R Liberal, and J W Pinney. Sbmlschematic: a tool for effective visualization of sbml models. *Bioinformatics*, submitted, 2015.

[102] N.D Price, J.A Papin, C. H. Schilling, and B. O Palsson. Genome-scale microbial in silico models: the constraints-based approach. *TRENDSin Biotechnology*, 21(4), 2003.

[103] K D Pruitt, G R Brown, S M Hiatt, F Thibaud-Nissen, A Astashyn, O Ermolaeva, C M Farrell, J Hart, M J Landrum, K M McGarvey, MR Murphy, N A O'Leary, S Pujar, B Rajput, S H Rangwala, L D Riddick, A Shkeda, H Sun, P Tamez, R E Tully, C Wallin, D Webb, J Weber, W Wu, M Dicuccio, P Kitts, D R Maglott, T D Murphy, and J M Ostell. Refseq: an update on mammalian reference sequences. *Nucleic Acids Res.*, 42, 2013.

[104] Overbeek R. The ergo genome analysis and discovery system. *Nucleic Acids Research*, 31(1):164–171, 2003.

[105] G. Rastogi, A. Bhalla, A. Adhikari, K.M. Bischoff, S.R. Hughes, L.P. Christopher, and R.K. Sani. Characterisation of thermostable cellulases producedby bacillus and geobacillus strains. *Bioresource Technology*, 101(22):8798–8806, 2010.

[106] E Raux, C Thermes, P Heathcote, A Rambach, and M J Warren. A role for salmonella typhimurium cbik in cobalamin (vitamin b12) and siroheme biosynthesis. *J. Bacteriol*, 179(10):3202–12, 1997.

[107] D A Rodionov, A H Vitreschak, A A Mironov, and M S Gelfand. Comparative genomics of the vitamin b12 metabolism and regulation in prokaryotes. *J Biol Chem*, 278(42):41148–59, 2003.

[108] N Saitou and Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4:406–425, 1987.

[109] E M Sampson, CL Johnson, and T A Bobik. Biochemical evidence that the pdus gene encodes a bifunctional cobalamin reductase. *Microbiology*, 151(4):1169–77, 2005.

[110] D J Sander and J J Joung. Crispr-cas systems for editing, regulating and targeting genomes. *Nat. Biotech*, 32:347355, 2014.

[111] J D Sander and J K Joung. Crispr-cas systems for editing, regulating and targeting genomes. *Nature Biotech*, 32:347–355, 2014.

[112] M Scheer, A Grote, A Chang, I Schomburg, C Munaretto, M Rother, C Shngen, M Stelzer, J Thiele, and D Schomburg. Brenda, the enzyme information system in 2011. *Nucleic Acids Research*, 39:D670–D710, 2011.

[113] I Schomburg, A Chang, and Schomburg D. *brenda*,enzyme data and metabolic information. *Nucleic Acid Res*, 30(1):47–49, 2002.

[114] I Schomburg, A Chang, C Ebeling, M Gremse, C Heldt, G Huhn, and Schomburg D. Brenda, the enzyme database: updates and major new developments. *Nucleic Acid Res*, 1(32), 2004.

[115] D Segre, D Vitkup, and G M Church. Analysis of optimality in natural and perturbed metabolic networks. *PNAS*, 99(23):151127, 2002.

[116] P Shannon, A Markiel, O Ozier, N S Baliga, J T Wang, D Ramage, N Amin, B Schwikowski, and T Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13(11):2498–504, 2003.

[117] S Shulami, O Gat, A L SONENSHEIN, and Y SHOHAM. The glucuronic acid utilization gene cluster from *Bacillus stearothermophilus T-6*. *J. Bacter.*, 181(12):36953704, 1960.

[118] K Smallbone and P Mendes. The genus *Geobacillus*. *Bacillus Genetic Stock Center,Catalog of Strains*, 3(7), 2001.

[119] K Smallbone and P Mendes. Large-scale metabolic models: From reconstruction to differential equations. *Indust Biotech*, 9(4):179–184, 2013.

[120] J D Smiley and G Ashwell. Uronic acid metabolism in bacteria. iii. purification and properties of d-altronic acid and d-mannonic acid dehydrases in *Escherichia coli*. *J. Biol. Chem.*, 235(6):1571–1575, 1960.

[121] C Smolke. *The Metabolic Pathway Engineering Handbook: Fundamentals*. CRC Press, Boca Raton, FL, 2010.

[122] H Song and S Y Lee. Production of succinic acid by bacterial fermentation. *Enz. Microb. Tech.*, 39:352–361, 2006.

[123] Bieniasz-Krzywiec Sroka, J., S. L., Gwozdz, D. Leniowski, J. Lacki, M. Markowski, C. Avignone-Rossa, M. E. Bushell, J. McFadden, and A. M. Kierzak. Acorn: A grid computing system for constraint based modeling and visualisation of the genome scale metabolic reaction networks via a web interface. *BMC Bioinformatics*, 12(196), 2011.

[124] J Stelling. Mathematical models in microbial systems biology. *Current opinion in Microbiology*, 7:513–518, 2004.

[125] P Stenmark, P Kursula, S Flodin, S Grslund, R Landry, P Nordlund, and H Schler. Crystal structure of human inosine triphosphatase. substrate binding and implication

of the inosine triphosphatase deficiency mutation p32t. *J Biol Chem*, 282(5):3182–7, 2007.

[126] D J Studholme. Some (bacilli) like it hot: genomics of geobacillus species. *Microb Biotechnol*, 8(1):40–48, 2015.

[127] H Suzuki, K Yoshida, and R Ohshimac. Polysaccharide-degrading thermophiles generated by heterologous gene expression in *Geobacillus kaustophilus* hta426. *Appl Environ Microbiol.*, 79(17), 2013.

[128] M Tamamoto, T Ikeda, M Ishii, and Y Igarashi. Carboxylation reaction catalyzed by 2-oxoglutarate:ferredoxin oxidoreductases from /textitHydrogenobacter thermophilus. *Extremophiles*, 14(1):79–85, 2010.

[129] K Tamura, G Stecher, D Peterson, A Filipski, and S Kumar. *mega*6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol and Evol*, 30:2725–2729, 2013.

[130] Y J Tang, R Sapra, D Joyner, T C Hazen, S Myers, D Reichmuth, H Blanch, and J D Keasling. Analysis of metabolic pathways and fluxes in a newly discovered thermophilic and ethanol-tolerant geobacillus strain. *Arch Microbiol*, 5(102):1377–86, 2009.

[131] A Tanner, L Bowater, S A Fairhurst, and S Bornemann. A closed conformation of *Bacillus subtilis* oxalate decarboxylase oxdc provides evidence for the true identity of the active site. *J Biol Chem*, 276:43627–34, 2001.

[132] T Tatusova, S Ciufo, B Fedorov, K O'Neill, and I Tolstoy. Refseq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Res.*, 42(1), 2014.

[133] Inc The MathWorks. Matlab and statistics toolbox release 2012b, 2012.

[134] I. Thiele and B. O. Palsson. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocol*, 5(1), 2010.

[135] H Tjalsma, J van den Dolder, W J Meijer, G Venema, S Bron, and J M van Dijl. The plasmid-encoded signal peptidase sipp can functionally replace the major signal peptidases sips and sipt of *Bacillus subtilis*. *J Bacteriol.*, 181(8):2448–54, 1999.

[136] Nazina TN1, Tourova TP, Poltaraus AB, Novikova EV, Grigoryan AA, Ivanova AE, Lysenko AM, Petrunyaka VV, Osipov GA, Belyaev SS, and Ivanov MV. Taxonomic study of aerobic thermophilic bacilli: descriptions of geobacillus subterraneus gen. nov., sp.

nov. and geobacillus uzenensis sp. nov. from petroleum reservoirs and transfer of bacillus stearothermophilus, bacillus thermocatenulatus, bacillus thermoleovorans, bacillus kaustophilus, bacillus thermodenitrificans to geobacillus as the new combinations g. stearothermophilus, g. th. *IJSEM*, 51:433–446, 2001.

[137] New York United Nations. Framework convention on climate charge: Adoption of the paris agreement. 2015.

[138] New York United Nations. Transforming our world: the 2030 agenda for sustainable development. 2015.

[139] H van Bokhorst-van de Veen, H Xie, E Esveld, H Abee, T nad Mastwijk, and M Nierop Groot. Inactivation of chemical and heat-resistant spores of *Bacillus* and *Geobacillus* by nitrogen cold atmospheric plasma evokes distinct changes in morphology and integrity of spores. *Food Microbiol*, 45:26–33, 2015.

[140] L Van Melderen and M Saavedra De Bast. Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genetics*, 5(3), 2009.

[141] A Varma and B O Palsson. Metabolic flux balancing: Basic concepts, scientific and practical use. *Nature Biotech*, 12:994 – 998, 1994.

[142] S.F Velick. Glyceraldehyde-3-phosphate dehydrogenase from muscle. *Methods in Enzymology*, 1:401, 1955.

[143] C Ward. Application of metabolic flux and transcriptome analyses to understanding the physiology of engineered *Geobacillus thermoglucosidasius. Thesis submitted for the degree of Doctor of Philosophy*, page Imperial College London, 2014.

[144] B W Wortham, M A Oliveira, and C N Patel. Polyamines in bacteria: Pleiotropic effects yet specific mechanisms. *Adv in Exper Med and Biol*, 603:106–115, 2007.

[145] J Xue, C M Murrieta, D C Rule, and W K Miller. Exogenous or l-rhamnose-derived 1,2-propanediol is metabolized via a pdud-dependent pathway in *Listeria innocua. Appl. Environ. Microbiol.*, 74(23):7073–7079, 2008.

[146] Y Yang, D Gilbert, and S Kim. Annotation confidence score for genome annotation: a genome comparison approach. *Bioinformatics*, 26(1):22–29, 2010.

[147] C Zhang, B Ji, A Mardinoglu, J Nielsen, and Q Hua. Logical transformation of genome scale metabolic models for gene level applications and analysis. *Bioinformatics*, 31(14):2324–2331, 2015.

[148] Y Zhao, M P Caspersa, T Abeea, R J Siezena, and R Korta. Complete genome sequence of *Geobacillus thermoglucosidans* tno-09.020, a thermophilic sporeformer associated with a dairy-processing environment. *J Bacteriol*, 194(15), 2012.

[149] Y Zhou, Y Liang, K H Lynch, J J Dennis, and D S Wishart. Phast: A fast phage search tool. *Nucl. Acids Res.*, 39(2), 2011.

[150] D R Ziegler. The geobacillus paradox: why is a thermophilic bacterial genus so prevalent on a mesophilic planet? *Microbiology*, 160:1–11, 2014.

[151] H Zur, E Ruppin, and T Shlomi. imat: an integrative metabolic analysis tool. *Bioinformatics*, 26(24), 2010.

Appendices

Gene assignments for plasmids p11955-01/02

Plasmid	length(nt)	strand	Function	has assigned function	asserted in pathway
pGTH11955-01	1131	neg-	Transposase	1	0
pGTH11955-01	339	neg-	Hypothetical protein	0	0
pGTH11955-01	927	pos+	Transporter, drug/metabolite exporter family	1	1
pGTH11955-01	1350	pos+	Transposase	1	0
pGTH11955-01	123	pos+	unassigned	0	0
pGTH11955-01	1179	neg-	Chloramphenicol resistance protein	1	1
pGTH11955-01	915	pos+	Transcriptional regulators, LysR family	1	1
pGTH11955-01	114	neg-	unassigned	0	0
pGTH11955-01	1098	neg-	NAD-dependent oxidoreductase	1	0
pGTH11955-01	501	neg-	Acetyltransferase, GNAT family (EC 2.3.1.-)	1	0
pGTH11955-01	1281	neg-	5-aminolevulinic acid synthase (EC 2.3.1.37)	1	1
pGTH11955-01	978	neg-	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)	1	1
pGTH11955-01	129	neg-	unassigned	0	0
pGTH11955-01	453	pos+	unassigned	0	0
pGTH11955-01	288	pos+	unassigned	0	0
pGTH11955-01	1197	neg-	Plasmid replication initiation protein	1	0
pGTH11955-01	108	neg-	unassigned	0	0
pGTH11955-01	105	neg-	unassigned	0	0
pGTH11955-01	117	pos+	Transposase	1	0
pGTH11955-01	735	pos+	Transcriptional regulator, GntR family	1	1
pGTH11955-01	153	neg-	unassigned	0	0
pGTH11955-01	1548	pos+	(S)-2-hydroxy-acid oxidase chain D (EC 1.1.3.15)	1	1
pGTH11955-01	1371	pos+	(S)-2-hydroxy-acid oxidase subunit GlcF (EC 1.1.3.15)	1	1
pGTH11955-01	294	pos+	unassigned	0	0
pGTH11955-01	708	pos+	Transcriptional regulator, GntR family	1	1
pGTH11955-01	1110	pos+	Spermidine/putrescine-binding protein	1	1
pGTH11955-01	1095	pos+	ABC transporter ATP-binding protein	1	0
pGTH11955-01	930	pos+	Transporter	1	1
pGTH11955-01	792	pos+	Spermidine/putrescine transport system permease prc	1	1
pGTH11955-01	417	pos+	Hypothetical protein	0	0
pGTH11955-01	1203	pos+	Aminobutyraldehyde dehydrogenase (EC 1.2.1.19)	1	1
pGTH11955-01	216	pos+	unassigned	0	0
pGTH11955-01	399	pos+	Transcriptional regulator, GntR family	1	1
pGTH11955-01	621	pos+	Hypothetical membrane associated protein	0	0
pGTH11955-01	738	neg-	Transposase	1	0
pGTH11955-01	549	neg-	Transposase	1	0
pGTH11955-01	1089	neg-	Transposase	1	0
pGTH11955-01	636	pos+	Hypothetical membrane associated protein	0	0
pGTH11955-01	195	neg-	Transposase	1	0
pGTH11955-01	342	pos+	unassigned	0	0
pGTH11955-01	1167	neg-	Transposase	1	0
pGTH11955-01	195	neg-	Transposase	1	0
pGTH11955-01	150	pos+	unassigned	0	0
pGTH11955-01	285	neg-	unassigned	0	0
pGTH11955-01	795	neg-	Chromosome partitioning protein parA	1	1
pGTH11955-01	144	neg-	unassigned	0	0
pGTH11955-01	150	neg-	unassigned	0	0
pGTH11955-01	1410	pos+	Sigma-54-dependent transcriptional activator	1	0
pGTH11955-01	1128	pos+	Hypothetical membrane spanning protein	0	0
pGTH11955-01	1389	pos+	Hydantoinase/oxoprolinase family	1	0
pGTH11955-01	1314	pos+	Cytosine permease	1	1
pGTH11955-01	1095	pos+	Hypothetical membrane spanning protein	0	0
pGTH11955-01	1557	pos+	Hydantoinase/oxoprolinase family	1	0
pGTH11955-01	528	pos+	unassigned	0	0
pGTH11955-01	240	pos+	Hypothetical protein	0	0
pGTH11955-01	228	pos+	unassigned	0	0
pGTH11955-01	213	pos+	Hypothetical protein	0	0
pGTH11955-01	477	pos+	L-alanyl-D-glutamate peptidase (EC 3.4.-.-)	1	1
pGTH11955-01	90	pos+	unassigned	0	0
pGTH11955-01	162	neg-	unassigned	0	0
pGTH11955-01	126	pos+	unassigned	0	0
pGTH11955-01	162	pos+	unassigned	0	0
pGTH11955-01	567	pos+	DNA integration/recombination/inversion protein	1	0
pGTH11955-01	123	pos+	unassigned	0	0
pGTH11955-01	186	neg-	unassigned	0	0
pGTH11955-01	237	pos+	unassigned	0	0

Gene assignments for plasmids p11955-01/02

pGTH11955-01	1176	pos+	Plasmid replication initiation protein	1	0
pGTH11955-01	288	pos+	unassigned	0	0
pGTH11955-01	129	pos+	unassigned	0	0
pGTH11955-01	267	pos+	unassigned	0	0
pGTH11955-01	921	pos+	Catechol-2,3-dioxygenase (EC 1.13.11.2)	1	1
pGTH11955-01	204	pos+	4-oxalocrotonate tautomerase (EC 5.3.2.-)	1	0
pGTH11955-01	1272	pos+	Acyl-CoA dehydrogenase (EC 1.3.99.3)	1	1
pGTH11955-01	546	pos+	Flavin reductase family protein	1	0
pGTH11955-01	402	pos+	Flavodoxin reductase family protein	1	0
pGTH11955-01	1926	pos+	Transcriptional regulator	1	1
pGTH11955-01	90	neg-	unassigned	0	0
pGTH11955-01	837	pos+	2-oxopent-4-enoate hydratase (EC 4.2.1.80)	1	1
pGTH11955-01	882	pos+	Acetaldehyde dehydrogenase (EC 1.2.1.10)	1	1
pGTH11955-01	1026	pos+	4-hydroxy-2-oxovalerate aldolase (EC 4.1.3.39)	1	1
pGTH11955-01	801	pos+	4-oxalocrotonate decarboxylase (EC 4.1.1.77)	1	0
pGTH11955-01	678	pos+	Carboxylesterase (EC 3.1.1.1)	1	0
pGTH11955-01	309	pos+	unassigned	0	0
pGTH11955-01	246	pos+	unassigned	0	0
pGTH11955-01	213	neg-	unassigned	0	0
pGTH11955-01	150	pos+	unassigned	0	0
pGTH11955-01	393	pos+	unassigned	0	0
pGTH11955-01	123	pos+	unassigned	0	0
pGTH11955-01	873	neg-	Hypothetical protein	0	0
pGTH11955-01	186	pos+	unassigned	0	0
pGTH11955-01	279	pos+	unassigned	0	0
pGTH11955-01	420	pos+	Death ON curing protein	1	0
pGTH11955-01	207	pos+	unassigned	0	0
pGTH11955-01	201	neg-	unassigned	0	0
pGTH11955-01	129	neg-	unassigned	0	0
pGTH11955-01	105	pos+	unassigned	0	0
pGTH11955-01	405	neg-	ImpB/MucB/SamB family protein	1	0
pGTH11955-01	303	neg-	ImpB/MucB/SamB family protein	1	0
pGTH11955-01	300	pos+	Hypothetical protein	0	0
pGTH11955-01	282	pos+	unassigned	0	0
pGTH11955-01	963	neg-	Transposase	1	0
pGTH11955-01	144	neg-	unassigned	0	0
pGTH11955-01	498	pos+	Hypothetical membrane spanning protein	0	0
pGTH11955-01	1485	pos+	Hypothetical membrane spanning protein	0	0
pGTH11955-01	153	pos+	unassigned	0	0
pGTH11955-01	1035	pos+	Bacitracin transport ATP-binding protein bcrA	1	0
pGTH11955-01	567	pos+	Bacitracin transport permease protein BCRB	1	0
pGTH11955-01	423	pos+	Bacitracin transport permease protein BCRB	1	0
pGTH11955-01	396	pos+	unassigned	0	0
pGTH11955-01	1167	neg-	Transposase	1	0
pGTH11955-01	1185	neg-	Transposase	1	0
pGTH11955-01	921	neg-	Integrase/recombinase (XerC/CodV family)	1	0
pGTH11955-01	258	neg-	Transposase	1	0
pGTH11955-02	921	pos+	Integrase/recombinase (XerC/CodV family)	1	0
pGTH11955-02	516	neg-	Glutamyl endopeptidase precursor (EC 3.4.21.19)	1	1
pGTH11955-02	255	neg-	Glutamyl endopeptidase precursor (EC 3.4.21.19)	1	1
pGTH11955-02	348	neg-	Hypothetical protein	0	0
pGTH11955-02	840	neg-	unassigned	0	0
pGTH11955-02	363	neg-	Hypothetical protein	0	0
pGTH11955-02	990	neg-	LtrC-like protein	1	0
pGTH11955-02	192	neg-	unassigned	0	0
pGTH11955-02	2118	neg-	DNA topoisomerase III (EC 5.99.1.2)	1	1
pGTH11955-02	2133	neg-	TraG/TraD family	1	0
pGTH11955-02	1935	neg-	NICKASE	1	0
pGTH11955-02	258	pos+	unassigned	0	0
pGTH11955-02	252	pos+	Hypothetical protein	0	0
pGTH11955-02	645	pos+	Hypothetical protein	0	0
pGTH11955-02	1149	pos+	Hypothetical protein	0	0
pGTH11955-02	387	pos+	unassigned	0	0
pGTH11955-02	477	neg-	RelB family protein	1	0
pGTH11955-02	1017	neg-	Hypothetical protein	0	0
pGTH11955-02	540	neg-	Hypothetical protein	0	0
pGTH11955-02	312	neg-	Hypothetical protein	0	0

pGTH11955-02	201	neg-	Hypothetical protein	0	0
pGTH11955-02	633	pos+	Transcriptional regulator, copG family	1	1
pGTH11955-02	243	pos+	unassigned	0	0
pGTH11955-02	1044	pos+	unassigned	0	0
pGTH11955-02	1851	pos+	TraE-like protein	1	0
pGTH11955-02	1032	pos+	Peptidoglycan-specific endopeptidase, M23 family	1	1
pGTH11955-02	546	pos+	unassigned	0	0
pGTH11955-02	1032	pos+	unassigned	0	0
pGTH11955-02	660	pos+	unassigned	0	0
pGTH11955-02	1515	pos+	Hypothetical protein	0	0
pGTH11955-02	501	neg-	Hypothetical protein	0	0
pGTH11955-02	228	pos+	unassigned	0	0
pGTH11955-02	246	pos+	unassigned	0	0
pGTH11955-02	180	pos+	unassigned	0	0
pGTH11955-02	636	pos+	Cell filamentation protein fic	1	0
pGTH11955-02	177	pos+	unassigned	0	0
pGTH11955-02	1068	pos+	Integrase/recombinase (XerC/CodV family)	1	0
pGTH11955-02	489	neg-	unassigned	0	0
pGTH11955-02	564	neg-	DNA repair protein radC	1	1
pGTH11955-02	138	neg-	unassigned	0	0
pGTH11955-02	354	neg-	unassigned	0	0
pGTH11955-02	588	neg-	Hypothetical protein	0	0
pGTH11955-02	396	pos+	unassigned	0	0
pGTH11955-02	120	neg-	Signal peptidase I (EC 3.4.21.89)	1	1
pGTH11955-02	240	neg-	Signal peptidase I (EC 3.4.21.89)	1	1
pGTH11955-02	189	pos+	Transposase	1	0
pGTH11955-02	213	pos+	Transposase	1	0
pGTH11955-02	543	neg-	Hypothetical protein	0	0
pGTH11955-02	501	pos+	Transposase	1	0
pGTH11955-02	675	pos+	Transposase	1	0
pGTH11955-02	828	pos+	Peptidase	1	0
pGTH11955-02	729	pos+	Hypothetical protein	0	0
pGTH11955-02	699	pos+	ABC transporter ATP-binding protein	1	0
pGTH11955-02	150	pos+	Transposase	1	0
pGTH11955-02	279	pos+	Transposase	1	0
pGTH11955-02	231	neg-	unassigned	0	0
pGTH11955-02	504	pos+	ATP-dependent DNA helicase (EC 3.6.1.-), uvrD-rep fam	1	1
pGTH11955-02	483	pos+	Hypothetical protein	0	0
pGTH11955-02	795	pos+	Integrase/recombinase (XerC/CodV family)	1	0

```
#scripts for PathwayBooster
"""
LICENSE

   PathwayBooster is an open-source software tool to support the
   comparison and curation of metabolic models.
   Copyright (C) 2012  Rodrigo Liberal, Beata K. Lisowska, David J. Leak,  John W. Pinney

   This program is free software: you can redistribute it and/or modify
   it under the terms of the GNU General Public License as published by
   the Free Software Foundation, either version 3 of the License, or
   (at your option) any later version.

   This program is distributed in the hope that it will be useful,
   but WITHOUT ANY WARRANTY; without even the implied warranty of
   MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the
   GNU General Public License for more details.

   You should have received a copy of the GNU General Public License
   along with this program.  If not, see <http://www.gnu.org/licenses/>

"""
def readBlast10ToFile(blast10FileTo):
   handleF=open(blast10FileTo)
   QueryDict={}
   QueryDict2={}
   lineL=handleF.readlines()
   for line in lineL:
     line=line.strip()
     lineS=line.split(",")
     queryId=lineS[0]

     if ":" in queryId:
       queryId= queryId.split(":")[1]
     if "_cdsid_" in queryId:
       queryId= queryId.split("_cdsid_")[1]
     lineS[0]=queryId
     queryId2=lineS[1]
     if ":" in queryId2:

       queryId2= queryId2.split(":")[1]
     if "_cdsid_" in queryId2:
       queryId2= queryId2.split("_cdsid_")[1]
     lineS[1]=queryId2
     if not(queryId in QueryDict):
        QueryDict[queryId]=[]
     flagUnique=True
     if len(QueryDict[queryId]) < 3:
       for coisas in QueryDict[queryId]:
         if queryId2 in coisas:
          flagUnique=False
       if flagUnique:
         QueryDict[queryId].append([queryId2,lineS[2],lineS[10],lineS[11]])
   return QueryDict
```

```
def readBlast10FromFile(blast10FileFrom):
   handleF=open(blast10FileFrom)
   QueryDict={}
   QueryDict2={}
   lineL=handleF.readlines()
   for line in lineL:
     line=line.strip()
     lineS=line.split(",")
     queryId=lineS[0]
     if ":" in queryId:
       queryId= queryId.split(":")[1]
     if "_cdsid_" in queryId:
       queryId= queryId.split("_cdsid_")[1]
     lineS[0]=queryId
     queryId2=lineS[1]
     if ":" in queryId2:
       queryId2= queryId2.split(":")[1]
     if "_cdsid_" in queryId2:
       queryId2= queryId2.split("_cdsid_")[1]
     if not(queryId in QueryDict):
        QueryDict[queryId]=[queryId2,lineS[2],lineS[10],lineS[11]]
   return QueryDict


def RecFirsHitBlastS(query1,query2):
  FirsHitDict1={}
  FirsHitDict2={}
  for ent in query1:
    firstEnt=query1[ent][0]
    if firstEnt in query2:
      if query2[firstEnt][0][0] == ent:
          FirsHitDict1[ent]=query1[ent]
          FirsHitDict2[firstEnt]=query2[firstEnt][0]

  return [FirsHitDict1,FirsHitDict2]

def RecNFirsHitBlastS(query1,query2,NHits=3):
  FirsHitDict1={}
  FirsHitDict2={}
  for ent1 in query1:
    if ent11[0] in query2:
     for ent2 in query2[ent11[0]][:NHits]:
       if ent2[0] == ent1:
          if not(ent1 in FirsHitDict1):
            FirsHitDict1[ent1]=[]
          if not(ent11 in FirsHitDict1[ent1]):
            FirsHitDict1[ent1].append(ent11)

  for ent2 in query2:
    for ent22 in query2[ent2][:NHits]:
     if ent22[0] in query1:
      for ent1 in query1[ent22[0]][:NHits]:
       if ent1[0] == ent2:
          if not(ent2 in FirsHitDict2):
            FirsHitDict2[ent2]=[]
          if not(ent22 in FirsHitDict2[ent2]):
            FirsHitDict2[ent2].append(ent22)
  return [FirsHitDict1,FirsHitDict2]
```

```
##########################
"""
LICENSE

   PathwayBooster is an open-source software tool to support the
   comparison and curation of metabolic models.
   Copyright (C) 2012  Rodrigo Liberal, Beata K. Lisowska, David J. Leak,  John W. Pinney

   This program is free software: you can redistribute it and/or modify
   it under the terms of the GNU General Public License as published by
   the Free Software Foundation, either version 3 of the License, or
   (at your option) any later version.

   This program is distributed in the hope that it will be useful,
   but WITHOUT ANY WARRANTY; without even the implied warranty of
   MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the
   GNU General Public License for more details.

   You should have received a copy of the GNU General Public License
   along with this program.  If not, see <http://www.gnu.org/licenses/>

"""
print
print   "PathwayBooster  Copyright (C) 2012"
print   "Rodrigo Liberal, Beata K. Lisowska, David J. Leak,  John W. Pinney"
print   "This program comes with ABSOLUTELY NO WARRANTY;"
print   "This is free software, and you are welcome to redistribute it under certain"# conditions;
see README.txt file for details."
print   "conditions; see README.txt file for details."
print
print

#class flushfile(object):
   #def __init__(self, f):
      #self.f = f
   #def write(self, x):
      #self.f.write(x)
      #self.f.flush()

#import sys

import traceback
try:
 flagsys=True
 try:
    import sys
    #sys.stdout = flushfile(sys.stdout)
 except:
    print "you need to install sys python package"
    flagsys=False
 import os
 doGeneral=True
 doBlast=True
 doHamming=True
 doDisplay=True
 doBrenda=True
 webB=True
 RunProgram=False
```

```
pathwayRepName="PathwayBoosterReport"
setupFile='ProtocolSetup.xml'
Makeblast=False
entryList=sys.argv
#print entryList
#print "22"+22
placeForOut=os.getcwd()
placePathBooster=''
helpF=False
if flagsys:
 if len(entryList)>0:
   shortpath = os.path.split(entryList[0])
   placePathBooster=shortpath[0]
 fg=True
 if len(entryList)>1:
  for argT in xrange(1,len(entryList)):
   if fg:
    if "-" in entryList[argT][0]:
      if "help" == entryList[argT][1:]:
        helpF=True
      if "blast" == entryList[argT][1:]:
          Makeblast=True
          RunProgram=True

      elif entryList[argT][1:] == "outDir":
       if len(entryList)>argT+1:
          pathwayRepName=entryList[argT+1]
          fg=False
       elif len(entryList) <= argT+1 or entryList[argT+1][0]=="-":
        print "argument after ", entryList[argT], " is missing"
        RunProgram=False
      else:
        print entryList[argT], " is not a valid argument"
        RunProgram=False
    else:

       setupFile=entryList[argT]
   else:
    fg=True

if os.path.isfile(setupFile):
   RunProgram=True
else:
   print "Setup file does not exist: ", setupFile
   RunProgram=False

if flagsys:
 if not(RunProgram) or len(entryList)==1 or helpF:
  print "[path/]PathwayBooster.py [setupFile] [-options]"
  print "options available:"
  print "\t -blast  - build the blast files"
  print "\t -outDir [OutputFolder] - save report in the OutputFolder folder"
  RunProgram=False

if RunProgram:
 pythonFiles=os.path.join(placePathBooster,"pythonFiles")
 sys.path.append( pythonFiles )
 import os
 #print os.environ
```

```
#print sys.version_info
print
print "Importing python packages"
print
try:
 import shutil
except:
 print "you need to install shutil python package"
 RunProgram=False
try:
 import os
except:
 print "you need to install os python package"
 RunProgram=False

try:
  import commands
except:
  print "you need to install commands python package"
  RunProgram=False
#try:
  #from SOAPpy import WSDL
  #wsdl = 'http://soap.genome.jp/KEGG.wsdl'
  #server = WSDL.Proxy(wsdl)
#except:
  #print "you need to install SOAPpy python package"
  #RunProgram=False
try:
  import PIL
except:
  print "you need to install PIL python package"
  RunProgram=False

try:
  import urllib
except:
  print "you need to install urllib python package"
  RunProgram=False
try:
  import urllib2
except:
  print "you need to install urllib2 python package"
  RunProgram=False

try:
  import matplotlib
  matplotlib.use('Agg')
  from matplotlib import pyplot as plt
  from matplotlib import cm as CM

except:
  print "warning: you need to install matplotlib python package. Hit map will no be done"
  doHamming=False
try:
  import webbrowser
except:
  print "you need to install webbrowser python package. Report will not open authomatically."
  webB=False
```

```
if RunProgram:
 stre="All python packages are installed"
 if not(doHamming):
   stre=stre+ " except matplotlib."
 else:
  stre=stre+"."
 print stre
sys.stdout.flush()

if RunProgram:
 try:
   from readFiles import *

 except:
   print "readFiles.py is missing or misplaced"
   RunProgram=False
   var = traceback.format_exc()
   print var
 try:
   from tools import *
 except:
   print "tools.py is missing or misplaced"
   RunProgram=False
 try:
   from blastRead import *

 except:
   print "blastRead.py is missing or misplaced"
   RunProgram=False
 try:
   from PNGcoloring import *
 except:
   print "PNGcoloring.py is missing or misplaced"
   RunProgram=False



sys.stdout.flush()
if RunProgram:

 setupFileInfo=readSetupFileXml(setupFile)
 #print setupFileInfo[4]

 sys.stdout.flush()
 if setupFileInfo == False:
   print setupFile, " file does not exist"
   RunProgram=False




sys.stdout.flush()
if RunProgram:

 [PathList,GeneralReportL,DisplayL,HammingL,blastL,BrendaL,IdsDictsAnotFile]=setupFileInfo

 DisplayL=DisplayL[:7]
 if PathList == [[]]:
   print "warning: No Pathways were selected"
```

```
    RunProgram = False
  if GeneralReportL == [[]]:
    print "warning: No Organisms were given"
    RunProgram = False
##   print blastL
##   print "22"+22
  if blastL == [[]]:
    blastL =[]
    print "warning: Blast Files is empty"
  sys.stdout.flush()
 if RunProgram:
  querySpc=GeneralReportL[0][0]

  blastSpeciesNameL=[]
  ppRest=[]
  if doBlast or Makeblast:
   isBlastDir=os.path.isdir("blastFiles")
   if not(isBlastDir):
     os.makedirs("blastFiles")

   BlastFilesList = os.listdir('blastFiles')
##    print "----------- Blast dir -----------"
##    print BlastFilesList
   ppRest=blastL[1:]

   for spc in blastL[1:]:
     blastSpeciesNameL.append(spc[0])

   dictBlastFiles={}
   print
   print "Checking Blast input files."
   blastSp=[]
   for gg in  xrange(len(blastL)):
    if blastL[gg][1]!="":
       if not(os.path.isfile(blastL[gg][1])):
         print
         print "Warning: The file "+blastL[gg][1]+" from the "+blastL[gg][0]+" species does not
exist."
         if gg == 0:
           doBlast=False
           Makeblast=False
           print
           print "Error in the query species Blast file. Blast will not be performed."
         blastL[gg][1]=""
       else:
         blastSp.append(blastL[gg][0])

####    print blastL
##    print 22+"22"
  if doBlast or Makeblast:
   print
   print "Checking previous built Blast files. If you want to redo the Blast files, please use the blast
option.",
   print

   if not(Makeblast):
    for bfiles in BlastFilesList:
     BlastFileLoc=os.path.join("blastFiles",bfiles)
```

```
     if bfiles != '' and bfiles != 'blast.log' and bfiles != "blastfile2.out" and bfiles !=
"blastfile1.out"and os.stat(BlastFileLoc).st_size >10000:
      flagPassToTrue=False
      flagPassFromTrue=False
      namel=bfiles.split("__")
##       print "aaaaaaaaaaaaaaaaaaaaaaaaaaaa"
##       print blastSp
##       print namel[0]
##       print namel
      if namel[1]==querySpc:
       if namel[0] in blastSp:
        if not(namel[0] in dictBlastFiles):
         dictBlastFiles[namel[0]]=[{},{},[],[]]
        if namel[2][0]=="T":
         BLDicts=readBlast10ToFile(BlastFileLoc)
         dictBlastFiles[namel[0]][1]=BLDicts
         flagPassToTrue=True
        if namel[2][0]=="F":
         BLDicts=readBlast10FromFile(BlastFileLoc)
         dictBlastFiles[namel[0]][0]=BLDicts
         flagPassFromTrue=True

   blastSpc=[]
   for sppc2 in ppRest:

     if not(sppc2[0] in dictBlastFiles):
         blastSpc.append([sppc2[1],sppc2[0]])
     elif dictBlastFiles[sppc2[0]][0] =={} or dictBlastFiles[sppc2[0]][1] =={}:
         blastSpc.append([sppc2[1],sppc2[0]])
   if blastSpc!=[]:
    blastlog=os.path.join("blastFiles","blast.log")
##     print "VVVVVVVVVVVVVVVVVVVVV"
##     print blastL[0][1]
##     print blastlog
    os.system("makeblastdb -in %s -dbtype prot > %s" % (blastL[0][1],blastlog))
    print
    print "Building Blast Files",
    sys.stdout.flush()
##     print blastSpc
    for spec in blastSpc:

     if spec[0] != '':
##         print "-------spec[0]"
##         print spec[0]
        os.system("makeblastdb -in %s -dbtype prot > %s" % (spec[0],blastlog))
    for spec in blastSpc:
##      print "-----spec------"
##      print spec
     if spec[0] != '':
      dictBlastFiles[spec[1]]=[{},{},[],[]]


      try:

       blastFile1=os.path.join("blastFiles","blastfile2.out")
       blastFile2=os.path.join("blastFiles","blastfile1.out")
       os.system("blastp -query %s -db %s -out %s -outfmt 10 > %s"
%(spec[0],blastL[0][1],blastFile2,blastlog))
```

```
        os.system("blastp -query %s -db %s -out %s -outfmt 10 > %s"
%(blastL[0][1],spec[0],blastFile1,blastlog))
       except:
        print "warning: problem with ",blastSpc[0][0], " or ",spec[0]," files"

       fileOutnameTo=spec[1]+"__"+querySpc+"__ToblastModel.out"
       fileOutnameFrom=spec[1]+"__"+querySpc+"__FromblastModel.out"
       blastFileLocBT=os.path.join("blastFiles",fileOutnameTo)
       blastFileLocBF=os.path.join("blastFiles",fileOutnameFrom)
       shutil.move(blastFile2,blastFileLocBT)
       shutil.move(blastFile1,blastFileLocBF)
       blastDictsT=readBlast10ToFile(blastFileLocBT)
       blastDictsF=readBlast10FromFile(blastFileLocBF)

       dictBlastFiles[spec[1]][1]=blastDictsT
       dictBlastFiles[spec[1]][0]=blastDictsF
  print "-- done"
  sys.stdout.flush()
  if not(Makeblast):

   GeneralReportSpc=[]
   if doGeneral:
    for spc in GeneralReportL:
     GeneralReportSpc.append(spc[0])

   HammingL
   HammingDistanceSpc=[]
   if doHamming:

    for spc in HammingL:
     HammingDistanceSpc.append(spc[0])

   speciesBrenda=BrendaL
   if doBrenda:
    brendaLoc=os.path.join(placePathBooster,"files","brenda_download.txt")
   try:
    BrendaDict=readBrenda(speciesBrenda,brendaLoc)
   except:
    doBrenda=False
    print "warning: brenda_download.txt file is missing"

   speciesDisplay=[]
   colorList=[]
   if doDisplay:

    from PIL import ImageFont
    from PIL import Image
    from PIL import ImageDraw
    for spc in DisplayL:
     speciesDisplay.append(spc[0])
     colorList.append(spc[1])
    for ccll in xrange(len(colorList)):
     colorFalse=True
     for ccll2 in xrange(len(colorList[ccll])):
      #coo = float(colorList[ccll][ccll2])
      #if coo < 1:
       #colorFalse=True

     colorList[ccll][ccll2]=int(colorList[ccll][ccll2])
```

```
 listColors=colorList

 speciesModel=speciesDisplay+filter(lambda x:x not in speciesDisplay,HammingDistanceSpc)
 speciesModel=speciesModel+filter(lambda x:x not in speciesModel,GeneralReportSpc)
 speciesModel=speciesModel+filter(lambda x:x not in speciesModel,blastSpeciesNameL)
 #print IdsDictsAnotFile
 #print "22"+33
 for AnnotationIds in IdsDictsAnotFile:

   if  IdsDictsAnotFile[AnnotationIds][0]!="kegg":
     EC2GoFileLoc= os.path.join(placePathBooster,"files","ec2go.txt")
     #print IdsDictsAnotFile[AnnotationIds]
     if os.path.isfile(IdsDictsAnotFile[AnnotationIds][1]):
      EmblFileGT=readErgoEmblFile2(IdsDictsAnotFile[AnnotationIds][1],EC2GoFileLoc)
      IdsDictsAnotFile[AnnotationIds].append(EmblFileGT)

     else: # EmblFileGT == False:
      print
      print "warning: annotation file from "+ IdsDictsAnotFile[AnnotationIds][0]+" species is
not valid - "+IdsDictsAnotFile[AnnotationIds][1]
       sys.stdout.flush()
 if RunProgram and not(Makeblast):

 if doBlast:

  for spcN in dictBlastFiles:
   dictBlastFiles[spcN][2]=RecFirsHitBlastS(dictBlastFiles[spcN][0],dictBlastFiles[spcN][1])

   dictBlastFiles[spcN][3]=dictBlastFiles[spcN][1]

   dictBlastFiles[spcN][0]=[]
   dictBlastFiles[spcN][1]=[]
 pathNamesFileLoc= os.path.join(placePathBooster,"files","pathwayNames.txt")
 [setsToPaths,pathNumbToName]=readPathwayFile(pathNamesFileLoc)
 OutPutdir=os.path.isdir(pathwayRepName)

 if not(OutPutdir):
   os.makedirs(pathwayRepName)
 tmpDirLoc= os.path.join(pathwayRepName,"tmpFile")
 tmpDirLocDir=os.path.isdir(tmpDirLoc)
 if not(tmpDirLocDir):
  os.makedirs(tmpDirLoc)




 logoFile=os.path.join(placePathBooster,"htmlFiles","blueishB.png")
 filePathGeneral="index.html"
 openInicial= os.path.join(pathwayRepName,filePathGeneral)
 shutil.copy(logoFile, pathwayRepName)
 handleFileInicialPage= open(openInicial,"w")
 handleFileInicialPage.write('<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01
Transitional//EN">\n')
 handleFileInicialPage.write('<html>\n')
 handleFileInicialPage.write('<body>\n')
 handleFileInicialPage.write('<br>')
 handleFileInicialPage.write('<br>')
 handleFileInicialPage.write('<IMG SRC="blueishB.png"  WIDTH=1100 HEIGHT=62>\n<BR>')
```

```
  handleFileInicialPage.write('<br>\n')
  handleFileInicialPage.write('<table border="0">\n')
  handleFileInicialPage.write('<tr>\n')
  handleFileInicialPage.write('<th><h1>Pathways</h1><th>\n')
  handleFileInicialPage.write('<th> </th>\n')
  handleFileInicialPage.write('<th> </th>\n')
  handleFileInicialPage.write('</tr>\n')
  isSet=False
  placePwd=''
  print
  print "Generating the Report (you might want to take a coffee):"
  print
  if '' in PathList:
     PathList.remove('')

  for setOrPath in PathList:

   if setOrPath != []:
    if type(setOrPath) == list:
       handleFileInicialPage.write('<tr>\n')
       handleFileInicialPage.write('<td><big><b>##</b></big></td><td><big><b>Other
Pathways</b></big></td>\n')
       handleFileInicialPage.write('</tr>\n')
       PathList2=setOrPath

       placePwd=os.path.join(pathwayRepName,"Other_Pathways")
       placeFromInicialP='Other_Pathways'
       othersPathDir=os.path.isdir(placePwd)
       if not(othersPathDir):
          os.makedirs(placePwd)
    elif setOrPath in setsToPaths:
       handleFileInicialPage.write('<tr>\n')

handleFileInicialPage.write('<td><big><b>'+setOrPath+'</b></big></td><td><big><b>'+setsTo
Paths[setOrPath][0]+'</b></big></td>\n')
       handleFileInicialPage.write('</tr>\n')
       PathList2=setsToPaths[setOrPath][1]
       placePwd=os.path.join(pathwayRepName,setOrPath)
       placeFromInicialP=setOrPath
       setPathDir=os.path.isdir(placePwd)
       if not(setPathDir):
          os.makedirs(placePwd)

    for paths in PathList2:


     handleFileInicialPage.write('<tr>\n<td> </td>\n')
     HamDistDict={}
     pathName=''
     if paths in pathNumbToName:
        pathName= pathNumbToName[paths]
     placePwdPath=os.path.join(placePwd,paths)
     PathwayDir=os.path.isdir(placePwdPath)
     if not(PathwayDir):
       os.makedirs(placePwdPath)
     NewRepLoc=os.path.join(placePwdPath,"setupFile.xml")
     shutil.copy(setupFile, NewRepLoc)

     inicialFilePlace=placeFromInicialP+"/"+paths+"/"+"DisplayTab.html"
```

```
    handleFileInicialPage.write('<td><a
href=\"'+inicialFilePlace+'\">'+paths+'</td><td>'+pathName+'</td>\n')
    handleFileInicialPage.write('</tr>\n')
##      print HammingDistanceSpc
    for spc in HammingDistanceSpc:
      HamDistDict[spc]=''
    print "paths->",paths
    sys.stdout.flush()
    if True:
     print "\tBuilding Pathway dictionary"#this may take a few minutes
     sys.stdout.flush()
     [PathEcNumbers,dictEc2Id,dictId2Ec,dictIdsSpecieskgml]=getSPathEcNumbersRest(paths)
     #print PathEcNumbers
     #print dictIdsSpecieskgml
     #print len(dictIdsSpecieskgml)
     #elems = server.get_elements_by_pathway('path:ec'+paths)
     elems=dictIdsSpecieskgml
     flagPathNotEmpty=True
     if len(elems) == 0:
        print "warning: pathway empty"
        sys.stdout.flush()
        flagPathNotEmpty=False
     if flagPathNotEmpty:
      ppaathdictIdsToEc=dictId2Ec
      ppaathdictIds=dictEc2Id
      #[ppaathdictIds,ppaathdictIdsToEc]=getpathwayEcsGenes(elems)


      annotIdTokeggId={}
      for AnnotationIds in IdsDictsAnotFile:
        #print IdsDictsAnotFile

        if  IdsDictsAnotFile[AnnotationIds][0]=="kegg":
         annotIdTokeggId[AnnotationIds]=IdsDictsAnotFile[AnnotationIds][1]
         newKeggDict={}

newKeggDict=getSSpecECGenes(PathEcNumbers,IdsDictsAnotFile[AnnotationIds][1],paths,dictI
dsSpecieskgml)
          #for ECNN in PathEcNumbers:

#newKeggDict[ECNN]=getSSpecECGenes(ECNN,IdsDictsAnotFile[AnnotationIds][1],server)
          IdsDictsAnotFile[AnnotationIds].append([[],newKeggDict,[],[]])

      HammingDictFinal={}
      DisplayDictFinal={}
      GenralDictFinal={}
      BlastDictFinal={}

      for ECNumb in PathEcNumbers:
       HammingDictFinal[ECNumb]={}
       DisplayDictFinal[ECNumb]={}
       GenralDictFinal[ECNumb]={}
       BlastDictFinal[ECNumb]={}

       for spcV in blastL:
         spcId=spcV[0]
         BlastDictFinal[ECNumb][spcId]={}
         for annots in spcV[2]:
          AnnotID=annots[0]
```

```
            #print IdsDictsAnotFile
            if ECNumb in IdsDictsAnotFile[AnnotID][-1][1]:
              for genesNN in IdsDictsAnotFile[AnnotID][-1][1][ECNumb]:
                if not(genesNN in BlastDictFinal[ECNumb][spcId]):
                  BlastDictFinal[ECNumb][spcId][genesNN]=[]
                BlastDictFinal[ECNumb][spcId][genesNN].append(AnnotID)
        for spcV in GeneralReportL:
          spcId=spcV[0]
          GenralDictFinal[ECNumb][spcId]={}
          for annots in spcV[1:]:
            AnnotID=annots[0]
            if ECNumb in IdsDictsAnotFile[AnnotID][-1][1]:
              for genesNN in IdsDictsAnotFile[AnnotID][-1][1][ECNumb]:
                if not(genesNN in GenralDictFinal[ECNumb][spcId]):
                  GenralDictFinal[ECNumb][spcId][genesNN]=[]
                GenralDictFinal[ECNumb][spcId][genesNN].append(AnnotID)



        for spcV in DisplayL:
          spcId=spcV[0]
          DisplayDictFinal[ECNumb][spcId]='0'
          hasBrenda=False
          if len(spcV[-1]) != 3:
            itermedspcV=spcV[2:-1]
            hasBrenda=True
          else:
            itermedspcV=spcV[2:]
          for annots in itermedspcV:
            AnnotID=annots[0]
            if ECNumb in IdsDictsAnotFile[AnnotID][-1][1]:
              if len(IdsDictsAnotFile[AnnotID][-1][1][ECNumb])>0:
                DisplayDictFinal[ECNumb][spcId]='1'
                break
            if hasBrenda and doBrenda:

              if ECNumb in BrendaDict and spcId in BrendaDict[ECNumb] and
len(BrendaDict[ECNumb][spcId]["in"])>0:
                DisplayDictFinal[ECNumb][spcId]='1'

        for spcV in HammingL:
          spcId=spcV[0]
          HammingDictFinal[ECNumb][spcId]='0'
          hasBrenda=False
          if len(spcV[-1]) != 3:
            itermedspcV=spcV[2:-1]
            hasBrenda=True
          else:
            itermedspcV=spcV[2:]
          for annots in itermedspcV:
            AnnotID=annots[0]
            if ECNumb in IdsDictsAnotFile[AnnotID][-1][1]:
              if len(IdsDictsAnotFile[AnnotID][-1][1][ECNumb])>0:
                HammingDictFinal[ECNumb][spcId]='1'
                break
            if hasBrenda and doBrenda: ##it means brenda is also considered
              if ECNumb in BrendaDict and spcId in BrendaDict[ECNumb] and
len(BrendaDict[ECNumb][spcId]["in"])>0:
```

```
                HammingDictFinal[ECNumb][spcId]='1'

        #print DisplayDictFinal
        #print "22"+22
        bkG=[]
        frG=[]
        idList=[]
        CheckPathEcNumbers=PathEcNumbers[0:]
        IdDictPresence={}
        for ecNumb in PathEcNumbers:

         #print ppaathdictIds
         #print "22"+22
         #if "ec:"+ecNumb in ppaathdictIds:
         #print PathEcNumbers
         #print ppaathdictIds
         #print ecNumb
         #print "22"+22
         if ecNumb in ppaathdictIds:

          for iddd in ppaathdictIds[ecNumb]:
            IdDictPresence[iddd]=''
            #print ppaathdictIds[ecNumb]
            #print iddd
            idList.append(iddd)
            binstring=''
            #binstring=binstring*(8-len(speciesDisplay))
            #frG.append("#303030")
            for specNN in speciesModel:
              if specNN in DisplayDictFinal[ecNumb]:
                 binstring=binstring+DisplayDictFinal[ecNumb][specNN]
              if specNN in HammingDictFinal[ecNumb]:
##                  print HamDistDict
                 HamDistDict[specNN]=HamDistDict[specNN]+HammingDictFinal[ecNumb][specNN]
            IdDictPresence[iddd]=binstring
            #decN=int(binstring,2)
            #hexN=hex(decN)[2:]
            #hexId2="00"
            #if iddd > 255:
                #hexId="FF"
                #hexId2=hex(iddd-255)[2:]
            #else:
              #hexId=hex(iddd)[2:]
            #if len(hexId) == 1:
              #hexId="0"+hexId
            #if len(hexId2) == 1:
              #hexId2="0"+hexId2
            #colour="#"+hexN+hexId+hexId2
            #bkG.append(colour)
          else:
            #print ppaathdictIds
            #print ecNumb
            #print "22"+22
            #print ecNumb
            print "warning: KEGG database may be updating."
            sys.stdout.flush()
            CheckPathEcNumbers.remove(ecNumb)
        #print IdDictPresence
        #print "22"+22
```

```
        print "\tReports part"
        PathEcNumbers=CheckPathEcNumbers
        if doDisplay:
           print "\t-pathway display"
           sys.stdout.flush()
           #linkStr=server.color_pathway_by_elements('path:ec'+paths, idList, frG, bkG)
           url = 'http://rest.kegg.jp/get/ec'+paths+'/image'
           response = urllib2.urlopen(url).read()
           fillettLoc=os.path.join(tmpDirLoc,'filettt.png')
           #print response
           fh = open(fillettLoc, "wb")
           fh.write(response)
           fh.close()

           #u = urllib.urlretrieve(linkStr,fillettLoc)

colorFile=putcolors(fillettLoc,listColors[:len(speciesDisplay)],tmpDirLoc,dictIdsSpecieskgml,IdDi
ctPresence)
           fontPath=os.path.join(placePathBooster,"files","LinBiolinum_R_G.ttf")
           placePathBooster
           font = ImageFont.truetype(fontPath,15)
           img=Image.new("RGBA", (colorFile[1]+5,colorFile[0]+100),(255,255,255))
           colorsLLeg=listColors[:len(speciesDisplay)]
           tti=[10,colorFile[0]+8]
           hhT=0
           draw = ImageDraw.Draw(img)
           wwidth=colorFile[1]
           longestStr=0

           for spcLe in speciesDisplay:
              if len(spcLe)>longestStr:
                 longestStr=len(spcLe)
           minimumAc=30+(9*longestStr)
           for spcLe in xrange(len(speciesDisplay)):
             if (wwidth - tti[0])<(minimumAc):
               tti[1]=tti[1]+31
               tti[0]=10
             draw.rectangle((tti[0], tti[1], tti[0]+15, tti[1]+15),
fill=(colorsLLeg[spcLe][0],colorsLLeg[spcLe][1],colorsLLeg[spcLe][2]))
             draw.text((tti[0]+20, tti[1]+1),speciesDisplay[spcLe],(0,0,0),font=font)
             tti[0]=tti[0]+minimumAc
             hhT=hhT+1
             draw = ImageDraw.Draw(img)
           draw = ImageDraw.Draw(img)
           draw = ImageDraw.Draw(img)
           atestpngLoc=os.path.join(tmpDirLoc,"a_test.png")
           ramppngLoc=colorFile[2]
           img.save(atestpngLoc)
           joinI=joinPngImages(atestpngLoc,ramppngLoc,tmpDirLoc)
           shutil.copy(joinI,placePwdPath)

htmlFile=makeHtmlPathFile(PathEcNumbers,paths,[joinI,[colorFile[1],colorFile[0]+100]],colorFi
le,ppaathdictIdsToEc,tmpDirLoc)
           shutil.copy(htmlFile,placePwdPath)
           os.remove(htmlFile)
           newHtmlDir=os.path.join(placePwdPath,"htmlFiles")
           isNewDir=os.path.isdir(newHtmlDir)
           if isNewDir:
              shutil.rmtree(newHtmlDir)
```

```
          HtmlOriginal=os.path.join(placePathBooster,"htmlFiles")
          shutil.copytree(HtmlOriginal,newHtmlDir)

        if doGeneral:
         print "\t-Gene Annotation report"
         sys.stdout.flush()
         generalLoc=os.path.join(placePwdPath,"GeneAnnotationsTab.html")
         handleFilePathRG= open(generalLoc,"w")
         handleFilePathRG.write('<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01
Transitional//EN">\n')
         handleFilePathRG.write('<html>\n')
         handleFilePathRG.write('<head>\n')
         commonHeadHtml(handleFilePathRG)
         handleFilePathRG.write('</head>\n')
         handleFilePathRG.write('<body>\n')
         commonBodyTopHtml(handleFilePathRG,2,paths)
         commonInicial2Html(handleFilePathRG)
         handleFilePathRG.write('<td><br></td></tr>\n')
         handleFilePathRG.write('</table>\n')
         handleFilePathRG.write('<table  class="sofT" border="0">\n')
         handleFilePathRG.write('<tr>\n')
         handleFilePathRG.write('<th class="helpHed" width="70">EC Number</th>\n')
         handleFilePathRG.write('<th class="helpHed" width="70">Species</th>\n')
         handleFilePathRG.write('<th class="helpHed" width="100">genes</th>\n')
         handleFilePathRG.write('<th class="helpHed" width="100">annotations</th>\n')
         handleFilePathRG.write('</tr>\n')
         for ECNumb in PathEcNumbers:
          ttt=0
          handleFilePathRG.write('<tr id="%s_%s_row">\n' %(ECNumb,str(ttt)))

          spcPresent=0
          NspanHtml=0
          ModelSpcPres="#7FFFD4"

          for spcNN in GeneralReportSpc:

             spcPAllD=GenralDictFinal[ECNumb][spcNN]
             if spcPAllD != {}:
                  spcPresent=spcPresent+1
                  NspanHtml=NspanHtml+len(spcPAllD)
             else:
                  if spcNN == GeneralReportSpc[0]:
                     ModelSpcPres='#D3D3D3'
                  NspanHtml=NspanHtml+1
          ttt=0
          handleFilePathRG.write('<td rowspan="%s" id="%s_1_cell"><a
name="%s_1"></a>%s</td>\n' %(str(NspanHtml),ECNumb,ECNumb,ECNumb))
          firstSpecies=True
          for spcNN in GeneralReportSpc:
            ttt=ttt+1
            if firstSpecies:

                handleFilePathRG.write('<td bgcolor=%s' %ModelSpcPres)
            else:
                handleFilePathRG.write('<tr id="%s_%s_row">\n' %(ECNumb,str(ttt)))
                handleFilePathRG.write('<td')
            firstSpecies=False
            handleFilePathRG.write(' rowspan="%s">%s</td>\n'
%(str(len(GenralDictFinal[ECNumb][spcNN])),spcNN))
```

```
          spcPrimeiro=True
          if len(GenralDictFinal[ECNumb][spcNN])==0:
            handleFilePathRG.write('</tr>\n')
            spcPrimeiro=False

          for gBs in GenralDictFinal[ECNumb][spcNN]:
           strHtmLink=gBs
           annotTypes=''
           for tAnnot in  GenralDictFinal[ECNumb][spcNN][gBs]:
              annotTypes=annotTypes+","+tAnnot
              if tAnnot in annotIdTokeggId:
                 strHtmLink='<a href="http://www.kegg.jp/dbget-
bin/www_bget?'+annotIdTokeggId[tAnnot]+':'+gBs+'">'+gBs+'</a>'
           annotTypes=annotTypes[1:]
           ttt=ttt+1
           if not(spcPrimeiro):

                handleFilePathRG.write('<tr id="%s_%s_row">\n' %(ECNumb,str(ttt)))
           spcPrimeiro=False
           handleFilePathRG.write('<td>%s</td>\n<td>%s</td>\n' %(strHtmLink,annotTypes))
           handleFilePathRG.write('</tr>\n')

        handleFilePathRG.write('</table>\n</div>\n')


        commonFinalHtml(handleFilePathRG)
        handleFilePathRG.write('</html>\n')


        handleFilePathRG.close()
       ########Hammilton distance part######


       if doHamming:
        print "\t-hamming distance"
        sys.stdout.flush()
        PathHDLoc=os.path.join(placePwdPath,"HeatMapTab.html")
        handleFilePathHD= open(PathHDLoc,"w")
        handleFilePathHD.write('<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01
Transitional//EN">\n')
        handleFilePathHD.write('<html>\n')
        handleFilePathHD.write('<head>\n')
        commonHeadHtml(handleFilePathHD)
        handleFilePathHD.write('</head>\n')
        handleFilePathHD.write('<body>\n')

        commonBodyTopHtml(handleFilePathHD,7,paths)
        handleFilePathHD.write('<div id="content">\n<br>\n')
        HamDistMatrix=hammiltonDistanceMatrix(HamDistDict,HammingDistanceSpc)
        nlistNames=[]
        for llkk in HammingDistanceSpc:
          nlistNames.append(llkk)

        ticksN=range(len(nlistNames))
        fig = plt.figure()
        plt.subplots_adjust(top=0.8,left=0.25)
        ax1 = fig.add_subplot(111)
```

```
        cmap = CM.get_cmap('RdBu', 100)
        cmap.set_bad('w')
        cax=ax1.matshow(HamDistMatrix,interpolation="nearest", cmap=cmap,vmin=0,
vmax=len(PathEcNumbers))
        ax1.set_xticks(ticksN, minor=False)
        ax1.set_yticks(ticksN, minor=False)
        for label in ax1.get_xticklabels():
            label.set_rotation(25)
            label.set_horizontalalignment('left')

        xtickNames = plt.setp(ax1, xticklabels=nlistNames,yticklabels=nlistNames)
        fig.colorbar(cax)
        ax1.grid(False)

        HamFigLoc=os.path.join(placePwdPath,"HeatMap.png")
        fig.savefig(HamFigLoc,dpi=100)
        handleFilePathHD.write('<img border="0" src="HeatMap.png" alt="Heat Map" width="620"
height="500" />')
        handleFilePathHD.write('</div>\n')
        handleFilePathHD.write('</html>\n')

      #######Blast search part#####


       if doBlast:
        print "\t-blast search"
        sys.stdout.flush()
        Hit1Loc=os.path.join(placePwdPath,"BlastBidirectionalHitTab.html")
        HitNLoc=os.path.join(placePwdPath,"Blast3BestHitsTab.html")
        handleFilePathRHit1=open(Hit1Loc,"w")
        handleFilePathRHitN=open(HitNLoc,"w")
        handleFilePathRHit1.write('<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01
Transitional//EN">\n')
        handleFilePathRHitN.write('<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01
Transitional//EN">\n')

        handleFilePathRHit1.write('<html>\n')
        handleFilePathRHit1.write('<head>\n')

        handleFilePathRHitN.write('<html>\n')
        handleFilePathRHitN.write('<head>\n')

        commonHeadHtml(handleFilePathRHit1)
        commonHeadHtml(handleFilePathRHitN)
        handleFilePathRHit1.write('</head>\n')
        handleFilePathRHit1.write('<body>\n')
        handleFilePathRHitN.write('</head>\n')
        handleFilePathRHitN.write('<body>\n')
        commonBodyTopHtml(handleFilePathRHit1,3,paths)
        commonBodyTopHtml(handleFilePathRHitN,4,paths)


        commonInicial2Html(handleFilePathRHit1)
        commonInicial2Html(handleFilePathRHitN)

        handleFilePathRHit1.write('<td><b>Query species - %s</b></td>\n' %querySpc)
        handleFilePathRHit1.write('</tr>\n')
        handleFilePathRHit1.write('</table>\n')
        handleFilePathRHitN.write('<td><b>Query species - %s</b></td>\n' %querySpc)
```

```
        handleFilePathRHitN.write('</tr>\n')
        handleFilePathRHitN.write('</table>\n')
        handleFilePathRHit1.write('<table class="sofT" border="0">\n')
        handleFilePathRHit1.write('<tr>\n')
        handleFilePathRHit1.write('<th class="helpHed" width="70">EC Number</th>\n')
        handleFilePathRHit1.write('<th class="helpHed" width="70">Target Species</th>\n')
        handleFilePathRHit1.write('<th class="helpHed" width="100">Target gene</th>\n')
        handleFilePathRHit1.write('<th class="helpHed" width="100">Query gene</th>\n')
        handleFilePathRHit1.write('<th class="helpHed" width="450">Query gene
function</th>\n')
        handleFilePathRHit1.write('<th class="helpHed" width="70">EC Number</th>\n')
        handleFilePathRHit1.write('<th class="helpHed" width="50">Seq. similarity</th>\n')
        handleFilePathRHit1.write('<th class="helpHed" width="50">e-value</th>\n')
        handleFilePathRHit1.write('<th class="helpHed" width="50">blast score</th>\n')
        handleFilePathRHit1.write('</tr>\n')


        handleFilePathRHitN.write('<table class="sofT" border="0">\n')
        handleFilePathRHitN.write('<tr>\n')
        handleFilePathRHitN.write('<th class="helpHed" width="70">EC Number</th>\n')
        handleFilePathRHitN.write('<th class="helpHed" width="70">Target Species</th>\n')
        handleFilePathRHitN.write('<th class="helpHed" width="100">Target gene</th>\n')
        handleFilePathRHitN.write('<th class="helpHed" width="100">Query gene</th>\n')
        handleFilePathRHitN.write('<th class="helpHed" width="450">Query gene
function</th>\n')
        handleFilePathRHitN.write('<th class="helpHed" width="70">EC Number</th>\n')
        handleFilePathRHitN.write('<th class="helpHed" width="50">Seq. similarity</th>\n')
        handleFilePathRHitN.write('<th class="helpHed" width="50">e-value</th>\n')
        handleFilePathRHitN.write('<th class="helpHed" width="50">blast score</th>\n')
        handleFilePathRHitN.write('</tr>\n')

        for Ecn in PathEcNumbers:

        ttt=0
        handleFilePathRHit1.write('<tr id="%s_%s_row" >\n' %(Ecn,str(ttt)))
        handleFilePathRHitN.write('<tr id="%s_%s_row" >\n' %(Ecn,str(ttt)))
        numberofSpc=0
        numberofSpcN=0
        dictSpcHtmlRows={}
        dictNSpcHtmlRows={}


        for scpN in blastSpeciesNameL:
          if scpN in dictBlastFiles:
            dictSpcHtmlRows[scpN]=0
            dictNSpcHtmlRows[scpN]=0
            passedinGen=False
            passeNdinGen=False
            for gen in BlastDictFinal[Ecn][scpN]:
              for ann in BlastDictFinal[Ecn][scpN][gen]:
                  TTT=gen
                  if TTT in IdsDictsAnotFile[ann][2][3]:
                    gen=IdsDictsAnotFile[ann][2][3][gen]
              if gen in dictBlastFiles[scpN][2][1]:
                passedinGen=True
                numberofSpc=numberofSpc+1
                dictSpcHtmlRows[scpN]=dictSpcHtmlRows[scpN]+1

              if gen in dictBlastFiles[scpN][3]:
```

```
                passeNdinGen=True
                if len(dictBlastFiles[scpN][3][gen])>0:
                   numberofSpcN=numberofSpcN+len(dictBlastFiles[scpN][3][gen])

dictNSpcHtmlRows[scpN]=dictNSpcHtmlRows[scpN]+len(dictBlastFiles[scpN][3][gen])
                else:
                   numberofSpcN=numberofSpcN+1
             if not(passedinGen):
                   numberofSpc=numberofSpc+1
                   dictSpcHtmlRows[scpN]=1
             if not(passeNdinGen):
                   numberofSpcN=numberofSpcN+1
                   dictNSpcHtmlRows[scpN]=1

          handleFilePathRHit1.write('<td rowspan="%s" id="%s_1_cell"><a
name="%s_1"></a>%s</td>\n' %(numberofSpc,Ecn,Ecn,Ecn))
          handleFilePathRHitN.write('<td rowspan="%s" id="%s_1_cell"><a
name="%s_1"></a>%s</td>\n' %(numberofSpcN,Ecn,Ecn,Ecn))

          firstSpcFlag=False
          firstSpcFlagN=False
          for scpN in blastSpeciesNameL:
           if scpN in dictBlastFiles:
            ttt=ttt+1
            if firstSpcFlag:
               handleFilePathRHit1.write('<tr id="%s_%s_row" >\n' %(Ecn,str(ttt)))
            firstSpcFlag=True
            handleFilePathRHit1.write('<td rowspan="%s">%s</td>\n'
%(str(dictSpcHtmlRows[scpN]),scpN))
            if firstSpcFlagN:
               handleFilePathRHitN.write('<tr id="%s_%s_row" >\n' %(Ecn,str(ttt)))
            firstSpcFlagN=True
            handleFilePathRHitN.write('<td rowspan="%s">%s</td>\n'
%(str(dictNSpcHtmlRows[scpN]),scpN))

            passedinGeneFlag=False
            passedinGeneFlagFirst=False
            passedinGeneFlagN=False
            passedinGeneFlagFirstN=False
            for gen in BlastDictFinal[Ecn][scpN]:
             prot=gen
             for ann in BlastDictFinal[Ecn][scpN][gen]:
                if prot in IdsDictsAnotFile[ann][2][3]:
                  prot=IdsDictsAnotFile[ann][2][3][gen]
             ProtGeneStr=''
             if prot!=gen:
                ProtGeneStr="("+prot+")"
             if prot in dictBlastFiles[scpN][2][1]:

                ttt=ttt+1
                if passedinGeneFlagFirst:
                      handleFilePathRHit1.write('<tr id="%s_%s_row">\n' %(Ecn,str(ttt)))
                passedinGeneFlagFirst=True
                ErgoGene=dictBlastFiles[scpN][2][1][prot][0]
                passedinGeneFlag=True
                gGfunction=''
                ECNNLL=[]
                FucntLL=[]
```

```
            Gen2=ErgoGene
            for typpes in IdsDictsAnotFile:


              if ErgoGene in IdsDictsAnotFile[typpes][2][2]:
                  Gen2=IdsDictsAnotFile[typpes][2][2][ErgoGene]

              if Gen2 in IdsDictsAnotFile[typpes][2][0]:
                  FucntLL.append(IdsDictsAnotFile[typpes][2][0][Gen2]["function"]+"-"+typpes)
                  gGfunction=gGfunction+" ##
"+IdsDictsAnotFile[typpes][2][0][Gen2]["function"]+"-"+typpes
                  for Tec in  IdsDictsAnotFile[typpes][2][0][Gen2]["EC"]:
                      if not(Tec in ECNNLL):
                          ECNNLL.append(Tec)
            if len(gGfunction) > 0 and gGfunction[0]== "/":
              gGfunction = gGfunction[1:]
            GeneProtStr2=''
            if Gen2!=ErgoGene:
               GeneProtStr2="("+ErgoGene+")"
            handleFilePathRHit1.write('<td>%s %s</td>\n<td>%s %s</td>\n<td>%s</td>\n'
%(gen,ProtGeneStr,Gen2,GeneProtStr2,gGfunction))
            handleFilePathRHit1.write('<td>\n<table class="subTable">\n')
            for ECNN in ECNNLL:
                ttt=ttt+1
                handleFilePathRHit1.write('<tr>\n<td>%s</td>\n</tr>\n' %(ECNN))
            handleFilePathRHit1.write('</table>\n</td>\n')
            handleFilePathRHit1.write('<td>%s</td>\n<td>%s</td>\n<td>%s</td>\n'
%(str(dictBlastFiles[scpN][2][1][prot][1]),str(dictBlastFiles[scpN][2][1][prot][2]),str(dictBlastFi
les[scpN][2][1][prot][3])))
            handleFilePathRHit1.write('</tr>\n')
          if prot in dictBlastFiles[scpN][3]:
            ttt=ttt+1
            if passedinGeneFlagFirstN:
                handleFilePathRHitN.write('<tr id="%s_%s_row">\n' %(Ecn,str(ttt)))
            passedinGeneFlagFirstN=True
            passedinGeneFlagN=True
            geneLen=1
            if len(dictBlastFiles[scpN][3][prot])>0:
              geneLen=len(dictBlastFiles[scpN][3][prot])
              handleFilePathRHitN.write('<td rowspan="%s">%s %s</td>\n'
%(geneLen,gen,ProtGeneStr))
            else:
              handleFilePathRHitN.write('<td>%s</td>\n</tr>\n' %prot)
            flagModelGen=False
            for hitsGene in dictBlastFiles[scpN][3][prot]:
              gGfunctionN=''
              ECNNLL=[]
              FucntLL=[]
              for typpes in IdsDictsAnotFile:
                Gen2=hitsGene[0]
                if hitsGene[0] in IdsDictsAnotFile[typpes][2][2]:
                   Gen2=IdsDictsAnotFile[typpes][2][2][hitsGene[0]]
                GeneProtStr2=''
                if Gen2!=hitsGene[0]:
                   GeneProtStr2="("+hitsGene[0]+")"
                if Gen2 in IdsDictsAnotFile[typpes][2][0]:
                  FucntLL.append(IdsDictsAnotFile[typpes][2][0][Gen2]["function"])
                  gGfunctionN=gGfunctionN+" ##
"+IdsDictsAnotFile[typpes][2][0][Gen2]["function"]+"-"+typpes
```

```
                for Tec in  IdsDictsAnotFile[typpes][2][0][Gen2]["EC"]:
                    if not(Tec in ECNNLL):
                        ECNNLL.append(Tec)
                gGfunctionN = gGfunctionN#[1:]
            ttt=ttt+1
            if flagModelGen:
                handleFilePathRHitN.write('<tr id="%s_%s_row">\n' %(Ecn,str(ttt)))
            flagModelGen=True
            handleFilePathRHitN.write('<td>%s %s</td>\n<td>%s</td>\n'
%(Gen2,GeneProtStr2,gGfunctionN))
            gGfunctionN=''
            handleFilePathRHitN.write('<td>\n<table class="subTable">\n')
            for ECNN in ECNNLL:
                ttt=ttt+1
                handleFilePathRHitN.write('<tr>\n<td>%s</td>\n</tr>\n'  %(ECNN))
            handleFilePathRHitN.write('</table>\n</td>\n')
            handleFilePathRHitN.write('<td>%s</td>\n<td>%s</td>\n<td>%s</td>\n'
%(str(hitsGene[1]),str(hitsGene[2]),str(hitsGene[3])))
            handleFilePathRHitN.write('</tr>\n')
        if not(passedinGeneFlag):
            handleFilePathRHit1.write('</tr>\n')
        if not(passedinGeneFlagN):
            handleFilePathRHitN.write('</tr>\n')
    handleFilePathRHit1.write('</table>\n</div>\n')
    handleFilePathRHitN.write('</table>\n</div>\n')
    commonFinalHtml(handleFilePathRHit1)
    commonFinalHtml(handleFilePathRHitN)
    handleFilePathRHit1.write('</html>\n')
    handleFilePathRHitN.write('</html>\n')
    handleFilePathRHit1.close()
    handleFilePathRHitN.close()

    if doBrenda:


    print "\t-publications"
    sys.stdout.flush()
    PublicationsLoc=os.path.join(placePwdPath,"PublicationsPositiveTab.html")
    PublicationsNotLoc=os.path.join(placePwdPath,"PublicationsNegativeTab.html")
    handlePublications=open(PublicationsLoc,"w")
    handlePublicationsNot=open(PublicationsNotLoc,"w")
    handlePublications.write('<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01
Transitional//EN">\n')
    handlePublicationsNot.write('<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01
Transitional//EN">\n')

    handlePublications.write('<html>\n')
    handlePublications.write('<head>\n')
    handlePublicationsNot.write('<html>\n')
    handlePublicationsNot.write('<head>\n')

    commonHeadHtml(handlePublications)
    commonHeadHtml(handlePublicationsNot)

    handlePublications.write('</head>\n')
    handlePublications.write('<body>\n')
    handlePublicationsNot.write('</head>\n')
    handlePublicationsNot.write('<body>\n')
```

```
        commonBodyTopHtml(handlePublications,5,paths)
        commonBodyTopHtml(handlePublicationsNot,6,paths)

        commonInicial2Html(handlePublications)
        commonInicial2Html(handlePublicationsNot)
        handlePublications.write('<td><br></td></tr>\n')
        handlePublications.write('</table>\n')

        handlePublications.write('<table class="sofT" border="0">\n')
        handlePublications.write('<tr>\n')
        handlePublications.write('<th class="helpHed" width="70">EC Number</th>\n')
        handlePublications.write('<th class="helpHed" width="70">Species</th>\n')
        handlePublications.write('<th class="helpHed" width="800">papers</th>\n')
        handlePublications.write('<th class="helpHed" width="120">pubmed number</th>\n')
        handlePublications.write('</tr>\n')

        handlePublicationsNot.write('<td><br></td></tr>\n')
        handlePublicationsNot.write('</table>\n')
        handlePublicationsNot.write('<table class="sofT" border="0">\n')
        handlePublicationsNot.write('<tr>\n')
        handlePublicationsNot.write('<th class="helpHed" width="70">EC Number</th>\n')
        handlePublicationsNot.write('<th class="helpHed" width="70">Species</th>\n')
        handlePublicationsNot.write('<th class="helpHed" width="800">papers</th>\n')
        handlePublicationsNot.write('<th class="helpHed" width="120">pubmed number</th>\n')
        handlePublicationsNot.write('</tr>\n')

        for ECNP in PathEcNumbers:
         if ECNP in  BrendaDict:
          ttt=0

          handlePublications.write('<tr id="%s_%s_row">\n' %(ECNP,str(ttt)))
          handlePublicationsNot.write('<tr id="%s_%s_row">\n' %(ECNP,str(ttt)))
          pubsSpeccount=0
          for speciesN in BrendaDict[ECNP]:
             if len(BrendaDict[ECNP][speciesN]["in"])>0:
                pubsSpeccount=pubsSpeccount+len(BrendaDict[ECNP][speciesN]["in"])
          if pubsSpeccount == 0:
            pubsSpeccount=1
          handlePublications.write('<td rowspan="%s" id="%s_1_cell"><a
name="%s_1"></a>%s</td>\n' %(pubsSpeccount,ECNP,ECNP,ECNP))
          pubNotsSpeccount=0
          for speciesN in BrendaDict[ECNP]:
             if len(BrendaDict[ECNP][speciesN]["out"])>0:
                pubNotsSpeccount=pubNotsSpeccount+len(BrendaDict[ECNP][speciesN]["out"])
          if pubNotsSpeccount==0:
            pubNotsSpeccount=1
          handlePublicationsNot.write('<td rowspan="%s" id="%s_1_cell"><a
name="%s_1"></a>%s</td>\n' %(pubNotsSpeccount,ECNP,ECNP,ECNP))
          pubpassedFirst=False
          pubNotpassedFirst=False
          for speciesN in BrendaDict[ECNP]:
           spcPub=1
           if len(BrendaDict[ECNP][speciesN]["in"])>0:
              spcPub=len(BrendaDict[ECNP][speciesN]["in"])
           if pubpassedFirst:
              ttt=ttt+1
              handlePublications.write('<tr id="%s_%s_row">\n' %(ECNP,str(ttt)))
           pubpassedFirst=True
           if len(BrendaDict[ECNP][speciesN]["in"])>0:
```

```
            spcPub=len(BrendaDict[ECNP][speciesN]["in"])
            handlePublications.write('<td rowspan="%s">%s</td>\n' %(spcPub,speciesN))
        pubpassedFirst2=False
        for pubS in BrendaDict[ECNP][speciesN]["in"]:
            ttt=ttt+1
            if pubpassedFirst2:
              handlePublications.write('<tr id="%s_%s_row">\n' %(ECNP,str(ttt)))
            pubpassedFirst2=True

            pubS=pubS.split("{")
            pubSN=pubS[-1].replace("}","")
            pubSN=pubSN.split(" ")[0]
            pubSN=pubSN.split("(")[0]
            pubSNN=pubSN.split(":")[1]
            if len(pubSNN)>3:
              strHtmLink='<a
href="http://www.ncbi.nlm.nih.gov/pubmed?term='+pubSNN+'">'+pubSN+'</a>'
            else:
              strHtmLink=  pubSN

            handlePublications.write('<td>%s</td>\n<td>%s</td>\n</tr>\n'
%(pubS[0],strHtmLink))
        if len(BrendaDict[ECNP][speciesN]["in"])==0:
            ttt=ttt+1
            handlePublications.write('</tr>\n')

        spcNotPub=1

        if pubNotpassedFirst:
            ttt=ttt+1
            handlePublicationsNot.write('<tr id="%s_%s_row">\n' %(ECNP,str(ttt)))
        pubNotpassedFirst=True
        if len(BrendaDict[ECNP][speciesN]["out"])>0:
            spcNotPub=len(BrendaDict[ECNP][speciesN]["out"])
            handlePublicationsNot.write('<td rowspan="%s">%s</td>\n'
%(spcNotPub,speciesN))
        pubNotpassedFirst2=False

        for pubS in BrendaDict[ECNP][speciesN]["out"]:
            ttt=ttt+1
            if pubNotpassedFirst2:
              handlePublicationsNot.write('<tr id="%s_%s_row">\n' %(ECNP,str(ttt)))
            pubNotpassedFirst2=True
            pubS=pubS.split("{")
            pubSN=pubS[-1].replace("}","")
            pubSN=pubSN.split(" ")[0]
            pubSNN=pubSN.split(":")[1]
            if len(pubSNN)>3:
              strHtmLink='<a
href="http://www.ncbi.nlm.nih.gov/pubmed?term='+pubSNN+'">'+pubSN+'</a>'
            else:
              strHtmLink=  pubSN

            handlePublicationsNot.write('<td>%s</td>\n<td>%s</td>\n</tr>\n'
%(pubS[0],strHtmLink))
        if len(BrendaDict[ECNP][speciesN]["out"])==0:
            handlePublicationsNot.write('</tr>\n')

    handlePublications.write('</table>\n</div>\n')
```

```
        handlePublicationsNot.write('</table>\n</div>\n')

        commonFinalHtml(handlePublications)
        commonFinalHtml(handlePublicationsNot)
        handlePublications.write('</html>\n')
        handlePublicationsNot.write('</html>\n')

        handlePublications.close()
        handlePublicationsNot.close()



    handleFileInicialPage.write('</table>\n')
    handleFileInicialPage.write('</body>\n')
    handleFileInicialPage.write('</html>\n')
    shutil.rmtree(tmpDirLoc)

    curentPath=os.getcwd()


    fileOpen=os.path.join("file:///",os.getcwd(),pathwayRepName,"index.html")
    fileOpen = "file:///"+fileOpen
    if webB:
      webbrowser.open(fileOpen)

except:
  RunProgram=False
  var = traceback.format_exc()
  print var

if RunProgram:
  print "finished"
else:
  print  "error - check warnings"
sys.stdout.flush()


################################################
```

```
####################

#script for python to parse SBML formated mets and get their Formula
A = open("reactions.xml", "r")
B = open("ReactionFormulaFull.txt", "w")
read_data = A.readlines()
import re
flag=0
for line in read_data:
        if line.strip().startswith("<reaction "):
                flag=1
                a=line.strip().split(" ")
                ID=a[1]
                Name=a[2]
                KEGG=a[3]
                B.write(Name + ':')
                #print Name, KEGG
        if line.strip().startswith("<listOfReactants>"):
                #print line
                flag=2
        if flag==2 and line.strip().startswith("<speciesReference "):
                b=line.strip().split(" ")
                b1=b[1]
                b2=b1.replace('species=', '')
                b3=b2.replace('"', '')
                Reactants=b3
                B.write( '\t' + Reactants )
                print Reactants
        if flag==2 and line.strip().startswith("</listOfReactants>"):
                #print line
                B.write('\t-->\t')
        if line.strip().startswith("<listOfProducts>"):
                #print line
                flag=4
        if flag==4 and line.strip().startswith("<speciesReference"):
                c=line.strip().split(" ")
                c1=c[1]
                c2=c1.replace('species=', '')
                c3=c2.replace('"', '')
                Product=c3
                print Product
                B.write(Product)
                #print Name, Reactants, '->', Product
                #B.write(Name + '\t' + KEGG + '\n' )
    if flag==4 and line.strip().startswith("</listOfProducts>"):
                #print line
                B.write('\n')


        #if line.strip().startswith("</reaction>")
        #       flag=FALSE
        #if flag and line.strip().startswith("<reaction ")

B.close()
####################

sbmlGeo = open("MetabolitesVersion7.txt", "r")
dictionary = open("MetabolitesVersion85.txt", "w")
read_data = sbmlGeo.readlines()
#read_data2 = sbmlBac.readlines()

import re
for line in read_data:
        re_findEC=re.compile("id=")
        findEC=re_findEC.search(line)
        if findEC:
                a=line.split(" ");
                b0=a[1]
                b1=b0.replace('"', '')
                b2=b1.split('=')
                b=b2[1]
                c=a[3]
```

```
                        d=a[5]
#                       print b, c, d
                        d2= d.replace('"', '')
                        e=d2.split("=")
                        bC=e[1]
                        #print b
                        if bC=="false":
                                c1=c.replace('"', '')
                                c2=c1.split("=")
                                C=c2[1]
                                #print C
                                if C=="Cytosol":
                                        b3=b+"_c"
                                        final=line.replace(b,b3)
                                        dictionary.write(final)
                                else:
                                        b4=b +"_e"
                                        final=line.replace(b,b4)
                                        print b4
                                        dictionary.write(final)

                        if bC=="true":
                                b5=b+"_b"
                                print b5
                                final=line.replace(b,b5)
                                print final
                        #print a
                                dictionary.write(final)

sbmlGeo.close()
dictionary.close()

####################

#script for python to parse SBML formatted mets and get their Formula
xmlA = open("mets2013_01_28.xml", "r")
f = open("metsWithFormula2013_01_28.xml", "w")
#dictionary1 = open("missingmetabolitesFINAL.txt", "w")
read_data = xmlA.readlines()
from xml.dom.minidom import Document

# Create the minidom document
doc = Document()

# Create the <xml> base element
xml = doc.createElement("sbml")
doc.appendChild(xml)

# Create the <model> child element to element sbml
mainmodel = doc.createElement("model")
mainmodel.setAttribute("metaid", "none")
xml.appendChild(mainmodel)
import re
#flag=0
for line in read_data:
        #flag=0
        re_findEC=re.compile("name=")
        findEC=re_findEC.search(line)
        if findEC:
                a=line.split('"')
                #print a
                id= a[1]
                name= a[3]
                compartment = a[5]
                charge =a[7]
                boundcond= a[9]
                constant =[11]
                b=a[3]
                compound=a[1]
                c=b.split('\t_')
                formula=c[1]
                specieslist = doc.createElement("ListofSpecies")
                mainmodel.appendChild(specieslist)
                #define species
```

```
                        species = doc.createElement("species")
                        species.setAttribute("id",id)
                        species.setAttribute("name",  name )
                        species.setAttribute("compartment",compartment)
                        species.setAttribute("charge", charge )
                        species.setAttribute("boundaryCondition",  boundcond )
                        species.setAttribute("constant","false")
                        specieslist.appendChild(species)

                        #set notes
                        notes=doc.createElement("notes")
                        species.appendChild(notes)

                        #set body

                        body=doc.createElement("body")
                        body.setAttribute("xmlns=", "http://www.w3.org/1999/xhtml")
                        notes.appendChild(body)
                        print boundcond
                        #set p
                        p=doc.createElement("p")
                        body.appendChild(p)

                        #put formula
                        ptext=doc.createTextNode("FORMULA: "+" "+ formula )
                        p.appendChild(ptext)

                        #put charge
                        p=doc.createElement("p")
                        body.appendChild(p)
                        ptext2=doc.createTextNode("CHARGE: "+" "+ charge )
                        p.appendChild(ptext2)


                        f.write(doc.toprettyxml(indent="  "))

f.close()
####################
B=open("GeobacillusEC+Genes_output.txt","w")
B.write("File created by Beata Lisowska\n")

B.write("\nGeobacillus thermoglucosidasius parsed EC numbers with RTMO genes which later will create dictionary for
GAGE Pathview\n\n")

#script for python to parse SBML formated mets and get their Formula
A = open("TMO_genbank.txt", "r")

read_data = A.readlines()
import re
flag=0
for line in read_data:
        if line.strip().startswith("CDS "):
                        flag=1
                        print "A"
        if flag==1 and line.strip().startswith("/locus_tag"):
                        a=line.strip().split(" ")
                        a=a[0]
                        a1=a.replace("/locus_tag=", "")
                        a2=a1.replace('"','')
                        print a2
                        flag=2
        if flag==2 and line.strip().startswith("/EC_number="):
                        b=line.strip().split(" ")
                        b=b[0]
                        b1=b.replace("/EC_number=","")
                        b2=b1.replace('"', '')
                        B.write(a2 + '\t' + 'ec:'+b2 )
                        flag=3
#                               if flag!=3 and flag==2:                         #
#                                       print "No EC number"

                        B.write('\n')

B.close()
```

```
###################
B=open("Dictionary.txt","w")
B.write("File created by Beata Lisowska\n")

B.write("\nGeobacillus thermoglucosidasius parsed EC numbers with  for GAGE Pathview\n\n")

A = open("GeobacillusEC+Genes_output.txt", "r")
C = open("C56ECtoGenes.txt", "r")
read_data1 = A.readlines()
read_data2 = C.readlines()

import re

for lineA in read_data1:
	if lineA.strip().startswith("RTMO"):
		a=lineA.strip().split("\t")
		RTMOgene=a[0]
		GtECnumber=a[1]
		for lineB in read_data2:
			b=lineB.strip(). split("\t")
			C56gene=b[0]
			C56ECnumber=b[1]
			if C56ECnumber==GtECnumber:
				B.write( RTMOgene + '\t' + C56gene +'\t'+ GtECnumber+'\t'+ C56ECnumber +
'\n')



A.close()
B.close()
C.close()
###################
A = open("/Users/b/Documents/PhD(2013-2014)/Models/RNAmodelAerobicSept2014.xml", "r+")
B = open("/Users/b/Documents/PhD(2013-2014)/CodeForFBA/RNAseqClear.txt", "r")
C = open("/Users/b/Documents/PhD(2013-2014)/CodeForFBA/AnaerobicModelRNAseqSeptember2014RPKM3.txt",
"w")
read_data1 = A.readlines()
read_data2 = B.readlines()
flag=0
import re
read_data2=read_data2[0].split("\r")
flagAllG=False
for line in read_data2:
	if line.startswith('RTM'):
		a=line.split('\t')
		gene1=a[0]
		RPKM1=a[1]
		RPKM2=a[2]
		RPKM3=a[3]
		RPKM4=a[4]
		allGene=[]
#		print gene1
		for lineB in read_data1:
		  #print(lineB)
			lineB=lineB.strip()
			findReaction=re.compile('<reaction')
			#print(findReaction)
			Reaction=findReaction.match(lineB)
			#print(Reaction)
			if Reaction:
				flagAllG='True
				allGene=[]
				b=lineB.split("")
				KEGGid=b[5]
				flag=1
			allGene.append(lineB)
			findGene=re.compile('<p>GENE_ASSOCIATION:')
			Gene=findGene.match(lineB)
			if Gene and flag==1:
				c=lineB.split(' ')
				gene2=c[2]
				flag=2
				if gene2==gene1 and flag==2:
					flag=3
```

```
                                findUB=re.compile('<parameter id="UPPER_BOUND"')
                                UB=findUB.match(lineB)
                                if UB and flag==3:
                                        d=lineB.split(' ')
                                        e=d[4].split('"')
                                        lineB=lineB.replace(e[1], RPKM3)
                                        flag=4
                                if flag==4 and flagAllG:
                                        flagAllG=False
                                        for kk in allGene:
                                                #print(kk)
                                        #print "C"
                                                C.write(kk + '\n')
                                        C.write(lineB +'\n')
                                        C.write('<parameter id="OBJECTIVE_COEFFICIENT" value="0"
units="mmol_per_gDW_per_hr" constant="false"/>' + '\n' + '<parameter id="FLUX_VALUE" value="0"
units="mmol_per_gDW_per_hr"
constant="false"/>'+'\n'+'</listOfParameters>'+'\n'+'</kineticLaw>'+'\n'+'</reaction>'+'\n')



A.close()
B.close()
C.close()
################
B=open("DictionaryC56+RTMO+Peg16Jan2015.txt","w")
#B.write("File created by Beata Lisowska\n")

#B.write("\nGeobacillus thermoglucosidasius with C56 genes for GAGE Pathview\n\n")

A = open("GeobacillusPegAndRTMO.txt", "r")
C = open("Dictionary.txt", "r")
read_data1 = A.readlines()
read_data2 = C.readlines()

import re

for lineA in read_data1:
        if lineA.strip().startswith("peg"):
                a=lineA.strip().split("\t")
                Peg=a[0]
                RTMOgene1=a[1]
                value1=a[2]
                value2=a[3]
                value3=a[4]
                value4=a[5]
                value5=a[6]
                value6=a[7]
                value7=a[8]
                value8=a[9]
                for lineB in read_data2:
#                       print lineB
                        b=lineB.strip(). split("\t")
                        RTMOgene2=b[0]
                        C56gene=b[1]
#                       print C56gene
                        if RTMOgene2==RTMOgene1:
                                B.write( Peg + '\t' + RTMOgene1 +'\t'+ C56gene + '\t'+ value1 +'\t'+ value2 +
'\t' +value3 + '\t'+ value4 + '\t'+ value5 +'\t'+ value6 + '\t' +value7 + '\t'+ value8 +'\n')

A.close()
B.close()
C.close()
```

Reactions with corresponding equation, EC number and FBA flux value under aerobic conditions in rich media

Name	Equation	Enzyme	KEGG RID	Flux aerobic
Pyrophosphate phosphohydrolase	|H2O| + |PPi| => (2) |Phosphate| + (2) |H+|	3.6.1.1	R00004	0
PHEt6	|L-Phenylalanine[e]| + |H+[e]| <=> |L-Phenylalanine| + |H+|	TC-2.A.3.1,2.A.3.1	None	27.9564
cytochrome oxidase bo3 (ubiquinol-8: 2.5 protons)	(0.5) |O2| + (2.5) |H+| + |Ubiquinol-8| => |H2O| + (2.5) |H+[e]| + |Ubiquinone-8|	Undetermined	None	1000
Cob(1)alamin transport via ABC system	|H2O| + |ATP| + |Cbl[e]| => |ADP| + |Phosphate| + |H+| + |Cbl|	Undetermined	None	0
isocitrate hydro-lyase	|Isocitrate| <=> |H2O| + |cis-Aconitate|	4.2.1.3	R01900	-1000
(2S,3R)-3-Hydroxybutane-1,2,3-tricarboxylate hydro-lyase	|Methylisocitrate| <=> |H2O| + |cis-2-Methylaconitate|	4.2.1.99	R04425	0
citrate hydro-lyase	|Citrate| <=> |H2O| + |cis-Aconitate|	4.2.1.3,4.2.1.4	R01325	1000
citrate hydro-lyase	|Citrate| <=> |Isocitrate|	4.2.1.3	R01324	-1000
glycerol transport in/out via diffusion reversible	|Glycerol| <=> |Glycerol[e]|	Undetermined	None	1000
ATP:glycerol 3-phosphotransferase	|ATP| + |Glycerol| <=> |ADP| + |Glycerol-3-phosphate|	2.7.1.30	R00847	733.926
N-Carbamoylputrescine amidohydrolase	|H2O| + (2) |H+| + |N-Carbamoylputrescine| => |CO2| + |NH3| + |Putrescine|	3.5.1.53	R01152	0
5,6-Dihydrouracil:NADP+ oxidoreductase	|NADP| + |Hydrouracil| <=> |NADPH| + |H+| + |Uracil|	1.3.1.2	R00978	0
5,6-Dihydrothymine:NADP+ oxidoreductase	|NADP| + |Dihydrothymine| <=> |NADPH| + |H+| + |Thymine|	1.3.1.2	R01415	0
5,6-Dihydrouracil amidohydrolase	|H2O| + |Hydrouracil| => |H+| + |3-Ureidopropanoate|	3.5.2.2	R02269	0
5,6-Dihydrothymine amidohydrolase	|H2O| + |Dihydrothymine| => |H+| + |3-Ureidoisobutyrate|	3.5.2.2	R03055	0
S-Adenosyl-L-methionine:8-amino-7-oxononanoate aminotransferase	|S-Adenosyl-L-methionine| + |8-Amino-7-oxononanoate| <=> |7-8-Diaminononanoate| + |S-Adenosyl-4	2.6.1.62	R03231	0
Succinyl-CoA:glycine C-succinyl-transferase(decarboxylating)	|Glycine| + |H+| + |Succinyl-CoA| => |CoA| + |CO2| + |5-Aminolevulinate|	2.3.1.37	R00830	0
Phosphoenolpyruvate:glycerone phosphotransferase	|Phosphoenolpyruvate| + |Glycerone| <=> |Pyruvate| + |Glycerone-phosphate|	2.7.1.121	R01012	-1000
Galactitol transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |Dulcose[e]| <=> |Pyruvate| + |Galactitol 1-phosphate|	Undetermined	None	0
Galactitol-1-phosphate:NAD oxidoreductase	|NAD| + |Galactitol 1-phosphate| <=> |NADH| + |H+| + |D-Tagatose 6-phosphate|	1.1.1.251	R05571	0
Acetoin dehydrogenase	|NAD| + |CoA| + |ACTN| <=> |NADH| + |Acetyl-CoA| + |H+| + |Acetaldehyde|	Undetermined	None	0
EX_pro_L_e	|L-Proline[e]| <=> |L-Proline|	Undetermined	None	-974.449
L-proline transport in via proton symport	|H+[e]| + |L-Proline[e]| <=> |H+| + |L-Proline|	Undetermined	None	1000
Hexosaminidase	|H2O| + |Chitobiose| => (2) |N-Acetyl-D-glucosamine|	3.2.1.52	R00022	0
4-Hydroxy-L-glutamate:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |4-Hydroxy-L-glutamate| <=> |L-Glutamate| + |4-Hydroxy-2-oxoglutarate|	2.6.1.23,2.6.1.1	R03266,R05052	0
L-Aspartate:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |L-Aspartate| <=> |L-Glutamate| + |Oxaloacetate|	2.6.1.1	R00355	955.99
sn-Glycerol ABC transport	|H2O| + |ATP| + |Glycerol-3-phosphate[e]| => |ADP| + |Phosphate| + |H+| + |Glycerol-3-phosphate|	Undetermined	None	863.762
Taurine, 2-oxoglutarate:O2 oxidoreductase (sulfite-forming)	|O2| + |2-Oxoglutarate| + |Taurine| => |CO2| + |Succinate| + |Sulfite| + |Aminoacetaldehyde|	1.14.11.17	R05320	0
D-Amino acid dehydrogenase	|H2O| + |FAD| + |D-Alanine| <=> |NH3| + |Pyruvate| + |FADH2|	1.4.99.1	None	-34.3887
TYRt6	|H+[e]| + |L-Tyrosine[e]| <=> |H+| + |L-Tyrosine|	TC-2.A.3.1,2.A.3.1	None	17.619
L-tryptophan transport in via proton symport	|L-Tryptophan[e]| + |H+[e]| <=> |L-Tryptophan| + |H+|	TC-2.A.3.1,2.A.3.1	None	8.64698
4a-hydroxytetrahydrobiopterin hydro-lyase	|4a-Hydroxytetrahydrobiopterin| <=> |H2O| + |Dihydrobiopterin|	4.2.1.96	R04734	0
Glycerol:NAD+ oxidoreductase	|NAD| + |Glycerol| <=> |NADH| + |H+| + |Glycerone|	1.1.1.6	R01034	-1000
(S)-2-Hydroxy-acid:oxygen 2-oxidoreductase	|O2| + |Glycolate| => |H2O2| + |Glyoxalate|	1.1.3.15	R00475	395.924
Glycolate oxidase	|Glycolate| + |Ubiquinone-8| <=> |Glyoxalate| + |Ubiquinol-8|	Undetermined	None	-1000
Glycolate oxidase	|Glycolate| + |Menaquinone 8| => |Glyoxalate| + |Menaquinol 8|	Undetermined	None	0
Glycolate oxidase	|Glycolate| + |2-Demethylmenaquinone 8| => |Glyoxalate| + |2-Demethylmenaquinol 8|	Undetermined	None	0
Propanoate:CoA ligase (ADP-forming)	|ATP| + |CoA| + |Propionate| <=> |ADP| + |Phosphate| + |Propionyl-CoA|	6.2.1.13	R00920	1000
Acetate:CoA ligase (AMP-forming)	|ATP| + |CoA| + |Acetate| + |H+| <=> |PPi| + |AMP| + |Acetyl-CoA|	6.2.1.1	R00235	-133.974
Propinol adenylate:CoA ligase (AMP-forming)	|CoA| + |Propionyladenylate| => |AMP| + |Propionyl-CoA|	6.2.1.1,6.2.1.17	R00926	0
Propanoate:CoA ligase (AMP-forming)	|ATP| + |H+| + |Propionate| <= |PPi| + |Propionyladenylate|	6.2.1.1,6.2.1.17	R01354	0
(S)-3-Hydroxybutanoyl-CoA:NAD+ oxidoreductase	|NAD| + |(S)-3-Hydroxybutyryl-CoA| <=> |NADH| + |H+| + |Acetoacetyl-CoA|	1.1.1.35,1.1.1.211	R01975	-1000
(2S,3S)-3-hydroxy-2-methylbutanoyl-CoA:NAD+ oxidoreductase	|NAD| + |2-methyl-3-hydroxy-butyryl-CoA| <=> |NADH| + |H+| + |2-Methylacetoacetyl-CoA|	1.1.1.35,1.1.1.178	R04203	0
(S)-3-Hydroxyhexadecanoyl-CoA:NAD+ oxidoreductase	|NAD| + |(S)-3-Hydroxyhexadecanoyl-CoA| <=> |NADH| + |H+| + |3-Oxopalmitoyl-CoA|	1.1.1.35,1.1.1.211	R04737	0
(S)-3-Hydroxydodecanoyl-CoA:NAD+ oxidoreductase	|NAD| + |(S)-3-Hydroxydodecanoyl-CoA| <=> |NADH| + |H+| + |3-Oxododecanoyl-CoA|	1.1.1.35,1.1.1.211	R04741	0
(S)-Hydroxydecanoyl-CoA:NAD+ oxidoreductase	|NAD| + |(S)-Hydroxydecanoyl-CoA| <=> |NADH| + |H+| + |3-Oxodecanoyl-CoA|	1.1.1.35,1.1.1.211	R04743	0
(S)-Hydroxyoctanoyl-CoA:NAD+ oxidoreductase	|NAD| + |(S)-Hydroxyoctanoyl-CoA| <=> |NADH| + |H+| + |3-Oxooctanoyl-CoA|	1.1.1.35,1.1.1.211	R04745	0
(S)-Hydroxyhexanoyl-CoA:NAD+ oxidoreductase	|NAD| + |(S)-Hydroxyhexanoyl-CoA| <=> |NADH| + |H+| + |3-Oxohexanoyl-CoA|	1.1.1.35,1.1.1.211	R04748	0
(S)-3-hydroxyacyl-CoA:NAD+ oxidoreductase	|NAD| + |(3S)-3-Hydroxyadipyl-CoA| <=> |NADH| + |H+| + |3-Oxoadipyl-CoA|	1.1.1.35	R06941	0
(S)-3-Hydroxytetradecanoyl-CoA:NAD+ oxidoreductase	|NAD| + |(S)-3-Hydroxytetradecanoyl-CoA| <=> |NADH| + |H+| + |3-Oxotetradecanoyl-CoA|	1.1.1.35,1.1.1.211	R04739	0
(S)-3-Hydroxybutanoyl-CoA:NADP+ oxidoreductase	|NADP| + |(S)-3-Hydroxybutyryl-CoA| <=> |NADPH| + |H+| + |Acetoacetyl-CoA|	1.1.1.157	R01976	1000
L-Glutamine-ABC transport	|H2O| + |ATP| + |L-Glutamine[e]| => |ADP| + |Phosphate| + |L-Glutamine| + |H+|	3.A.1.3	None	0
Glycerol:NAD+ oxidoreductase	|NAD| + |Glycerol| <=> |NADH| + |H+| + |D-Glyceraldehyde|	1.1.1.1,1.1.1.2,1.1.1.21,1	R01036	-729.838
Ethanol:NAD+ oxidoreductase	|NAD| + |Ethanol| <=> |NADH| + |H+| + |Acetaldehyde|	1.1.1.1,1.1.1.71	R00754	0
transport of (r)-3-hydroxybutanoate [extraorganism-cytosol](passive)	|(R)-3-Hydroxybutanoate[e]| <=> |(R)-3-Hydroxybutanoate|	Undetermined	None	0
3-Oxopropanoate:NAD+ oxidoreductase (decarboxylating,	|NAD| + |CoA| + |3-Oxopropanoate| <=> |NADH| + |CO2| + |Acetyl-CoA|	1.2.1.18,1.2.1.27	R00705	0
2-Methyl-3-oxopropanoate:NAD+ oxidoreductase (CoA-propanoylating)	|NAD| + |CoA| + |3-Oxo-2-methylpropanoate| <=> |NADH| + |CO2| + |Propionyl-CoA|	1.2.1.27	R00935,R00922	0
D-Mannitol-1-phosphate:NAD+ 5-oxidoreductase	|NAD| + |D-mannitol-1-phosphate| <=> |NADH| + |H+| + |D-fructose-6-phosphate|	1.1.1.17	R00758,R02703	1000
Glycerone phosphate phosphohydrolase (alkaline optimum)	|H2O| + |Glycerone-phosphate| => |Phosphate| + |H+| + |Glycerone|	3.1.3.1	R01010	0
4-Nitrophenyl phosphate phosphohydrolase	|H2O| + |4-Nitrophenyl phosphate| => |Phosphate| + |H+| + |PNP|	3.1.3.1,3.1.3.2,3.1.3.41	R03024	0
2-Amino-4-hydroxy-6-(erythro-1,2,3-trihydroxypropyl)	(3) |H2O| + |7,8-Dihydroneopterin 3'-triphosphate| => (3) |Phosphate| + (3) |H+| + |Dihydroneop	3.1.3.1	R04620	3.77943
ATP:D-Gluconate 6-phosphotransferase	|ATP| + |GLCN| <=> |ADP| + |6-Phospho-D-gluconate|	2.7.1.12	R01737	0
Urea amidohydrolase	|H2O| + (2) |H+| + |Urea| => |CO2| + (2) |NH3|	3.5.1.5	R00131	0
5-Aminopentanamide amidohydrolase	|H2O| + |5-Aminopentanamide| => |NH3| + |5-Aminopentanoate|	3.5.1.30	R02273	0
4-Aminobutanoate:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |GABA| <=> |L-Glutamate| + |4-Oxobutanoate|	2.6.1.19	R01648	0
L-ribulose-5-phosphate 4-epimerase	|L-ribulose-5-phosphate| <=> |D-Xylulose5-phosphate|	5.1.3.4	R05850	0
ATP:D-ribulose 5-phosphotransferase	|ATP| + |D-Ribulose| <=> |ADP| + |D-Ribulose5-phosphate|	2.7.1.16,2.7.1.47	R01526	0
ATP:L-ribulose 5-phosphotransferase	|ATP| + |L-Ribulose| <=> |ADP| + |L-ribulose-5-phosphate|	2.7.1.16	R02439	0
L-Arabinose ketol-isomerase	|L-Arabinose| <=> |L-Ribulose|	5.3.1.4	R01761	0
alpha-N-arabinofuranosidase	(2) |H2O| + |Arabinan| => (3) |L-Arabinose|	3.2.1.55	None	0
alpha-D-Glucose 1-epimerase	|D-Glucose| <=> |beta-D-Glucose|	5.1.3.3	R01602	-1000
xylose isomerase	|D-Glucose| <=> |D-Fructose|	5.3.1.5	R00878,R00307	-167.762
D-Xylose ketol-isomerase	|Xylose| <=> |D-Lyxulose|	5.3.1.5	R01432	0
ATP:D-xylulose 5-phosphotransferase	|ATP| + |D-Lyxulose| <=> |ADP| + |D-Xylulose5-phosphate|	2.7.1.17	R01639	0
cellobiose transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |CELB[e]| <=> |Pyruvate| + |cellobiose 6-phoshate|	Undetermined	None	832.238
beta-D-Glucoside glucohydrolase	|H2O| + |CELB| => (2) |beta-D-Glucose|	3.2.1.21	R00026	0
1,4-beta-D-Glucan glucohydrolase	|H2O| + |CELB| => (2) |D-Glucose|	3.2.1.74,3.2.1.21	R00306	0
beta-glucosidase (methyl-alpha-D-glucoside)	|H2O| + |alpha-Methyl-D-glucoside| <=> |D-Glucose| + |Methanol|	3.2.1.21	None	0
beta-glucosidase	|H2O| + |beta-Methylglucoside| <=> |D-Glucose| + |Methanol|	3.2.1.21	None	0
palmitoyl-UDP-glucosyltransferase (diglucosyl)	|UDP-glucose| + |Monoglucosyl-1,2 dipalmitoylglycerol| <=> |UDP| + |Diglucosyl-1,2 dipalmitoylglycer	Undetermined	None	0
myristoyl-UDP-glucosyltransferase (diglucosyl)	|UDP-glucose| + |Monoglucosyl-1,2 dimyristoylglycerol| <=> |UDP| + |Diglucosyl-1,2 dimyristoylglycer	Undetermined	None	0
isotetradecanoyl-UDP-glucosyltransferase (diglucosyl)	|UDP-glucose| + |Monoglucosyl-1,2 diisotetradecanoylglycerol| <=> |UDP| + |Diglucosyl-1,2 diisotetrac	Undetermined	None	0
isopentadecanoyl-UDP-glucosyltransferase (diglucosyl)	|UDP-glucose| + |Monoglucosyl-1,2 diisopentadecanoylglycerol| <=> |UDP| + |Diglucosyl-1,2 diisopent	Undetermined	None	0
anteisopentadecanoyl-UDP-glucosyltransferase (diglucosyl)	|UDP-glucose| + |Monoglucosyl-1,2 dianteisopentadecanoylglycerol| <=> |UDP| + |Diglucosyl-1,2 diant	Undetermined	None	0
isohexadecanoyl-UDP-glucosyltransferase (diglucosyl)	|UDP-glucose| + |Monoglucosyl-1,2 diisohexadecanoylglycerol| <=> |UDP| + |Diglucosyl-1,2 diisohexad	Undetermined	None	0
palmitoyl-UDP-glucosyltransferase (monoglucosyl)	|UDP-glucose| + |1,2-Diacyl-sn-glycerol dihexadecanoyl| <=> |UDP| + |Monoglucosyl-1,2 dipalmitoylgly	Undetermined	None	0
myristoyl-UDP-glucosyltransferase (monoglucosyl)	|UDP-glucose| + |1,2-Diacyl-sn-glycerol ditetradecanoyl| <=> |UDP| + |Monoglucosyl-1,2 dimyristoylgly	Undetermined	None	0
isotetradecanoyl-UDP-glucosyltransferase (monoglucosyl)	|UDP-glucose| + |1,2-Diisotetradecanoyl-sn-glycerol| <=> |UDP| + |Monoglucosyl-1,2 diisotetradecano	Undetermined	None	0
isopentadecanoyl-UDP-glucosyltransferase (monoglucosyl)	|UDP-glucose| + |1,2-Diisopentadecanoyl-sn-glycerol| <=> |UDP| + |Monoglucosyl-1,2 diisopentadecan	Undetermined	None	0
anteisopentadecanoyl-UDP-glucosyltransferase (monoglucosyl)	|UDP-glucose| + |1,2-Dianteisopentadecanoyl-sn-glycerol| <=> |UDP| + |Monoglucosyl-1,2 dianteisope	Undetermined	None	0
isohexadecanoyl-UDP-glucosyltransferase (monoglucosyl)	|UDP-glucose| + |1,2-Diisohexadecanoyl-sn-glycerol| <=> |UDP| + |Monoglucosyl-1,2 diisohexadecano	Undetermined	None	0
stearoyl-UDP-glucosyltransferase (diglucosyl)	|UDP-glucose| + |Monoglucosyl-1,2 distearoylglycerol| <=> |UDP| + |Diglucosyl-1,2 distearoylglycerol|	Undetermined	None	1.14629
isoheptadecanoyl-UDP-glucosyltransferase (diglucosyl)	|UDP-glucose| + |Monoglucosyl-1,2 diisoheptadecanoylglycerol| <=> |UDP| + |Diglucosyl-1,2 diisohept	Undetermined	None	1.14629
anteisoheptadecanoyl-UDP-glucosyltransferase (diglucosyl)	|UDP-glucose| + |Monoglucosyl-1,2 dianteisoheptadecanoylglycerol| <=> |UDP| + |Diglucosyl-1,2 diant	Undetermined	None	1.14629
stearoyl-UDP-glucosyltransferase (monoglucosyl)	|UDP-glucose| + |1,2-Diacyl-sn-glycerol dioctadecanoyl| <=> |UDP| + |Monoglucosyl-1,2 distearoylglyc	Undetermined	None	1.14629
isoheptadecanoyl-UDP-glucosyltransferase (monoglucosyl)	|UDP-glucose| + |1,2-Diisoheptadecanoyl-sn-glycerol| <=> |UDP| + |Monoglucosyl-1,2 diisoheptadecar	Undetermined	None	1.14629
anteisoheptadecanoyl-UDP-glucosyltransferase (monoglucosyl)	|UDP-glucose| + |1,2-Dianteisoheptadecanoyl-sn-glycerol| <=> |UDP| + |Monoglucosyl-1,2 dianteisohe	Undetermined	None	1.14629
ATP:D-fructose-1-phosphate 6-phosphotransferase	|ATP| + |D-fructose-1-phosphate| <=> |ADP| + |D-fructose-1,6-bisphosphate|	2.7.1.56	R02071	270.162
ATP:D-tagatose-6-phosphate 1-phosphotransferase	|ATP| + |D-Tagatose 6-phosphate| <=> |ADP| + |D-Tagatose 1,6-biphosphate|	2.7.1.11,2.7.1.144,2.7.1.:	R03236	0
D-fructose transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |D-Fructose[e]| <=> |Pyruvate| + |D-fructose-1-phosphate|	Undetermined	None	1000
Nitrate reductase (Menaquinol-8)	(2) |H+| + |Nitrate| + |Menaquinol 8| <=> |H2O| + (2) |H+[e]| + |Nitrite| + |Menaquinone 8|	Undetermined	None	-997.48
Nitrate reductase (Ubiquinol-8)	(2) |H+| + |Nitrate| + |Ubiquinol-8| => |H2O| + (2) |H+[e]| + |Nitrite| + |Ubiquinone-8|	1.7.99.4	None	0
Nicotinamide amidohydrolase	|H2O| + |Nicotinamide| => |NH3| + |Niacin|	3.5.1.19	R01268	0
Nicotinate D-ribonucleotide:pyrophosphate phosphoribosyltransferase	|PPi| + |Nicotinate ribonucleotide| <=> |H+| + |PRPP| + |Niacin|	2.4.2.11	R01724	-1.25981
sn-Glycero-3-phosphocholine glycerophosphohydrolase	|H2O| + |Glycerophosphocholine| => |Glycerol-3-phosphate| + |Choline|	3.1.4.2,3.1.4.46	R01030	0
sn-Glycero-3-phosphoethanolamine glycerophosphohydrolase	|H2O| + |Glycerophosphoethanolamine| => |Glycerol-3-phosphate| + |Aminoethanol|	3.1.4.2,3.1.4.46	R01470	0
Glycerophosphodiester phosphodiesterase (Glycerophosphoglycerol)	|H2O| + |Glycerophosphoglycerol| => |Glycerol-3-phosphate| + |Glycerol|	3.1.4.46	None	0
S-Adenosyl-L-methionine:precorrin-3B C17-methyltransferase	|S-Adenosyl-L-methionine| + |Precorrin 3B| <=> |S-Adenosyl-homocysteine| + (2) |H+| + |Precorrin 4|	2.1.1.131	R05180	0
precorrin-3B C17-methyltransferase	|S-Adenosyl-L-methionine| + |Cobalt-precorrin 3| <=> |S-Adenosyl-homocysteine| + |H+| + |Cobalt-pre	2.1.1.131	R05809	0
sirohydrochlorin cobalt-lyase	(2) |H+| + |Co2+| + |Sirohydrochlorin| <=> |Cobalt-precorrin 2|	4.99.1.3	R05807	0
precorrin-6Y:NADP+ oxidoreductase	|NADP| + |Precorrin 6B| <=> |NADPH| + |H+| + |Precorrin 6A|	1.3.1.54	R05150	0
precorrin-6A reductase	|NADPH| + |H+| + |Cobalt-precorrin 6| <=> |NADP| + |Cobalt-precorrin 6B|	1.3.1.54	R05812	0
Precorrin 8X 11,12-methylmutase	|H+| + |Precorrin 8| <=> |Hydrogenobyrinate|	5.4.1.2	R05177	0
precorrin-8X methylmutase	|Cobalt-precorrin 8| <=> |Cobyrinate|	5.4.1.2	R05814	0
rxn07587	|S-Adenosyl-L-methionine| + |Cobalt-precorrin 5B| <=> |S-Adenosyl-homocysteine| + |H+| + |Cobalt-p	Undetermined	R07773	0
S-Adenosyl-L-methionine:precorrin-4 C20-methyltransferase	|S-Adenosyl-L-methionine| + |H+| + |Precorrin 2| <=> |S-Adenosyl-homocysteine| + |Precorrin 3A|	2.1.1.130	R03948	0
S-adenosyl-L-methionine:cobalt-factor-II C20-methyltransferase	|S-Adenosyl-L-methionine| + |Cobalt-precorrin 2| <=> |S-Adenosyl-homocysteine| + |H+| + |Cobalt-pre	2.1.1.151	None	0

S-adenosyl-L-methionine:precorrin-4 C11 methyltransferase	|S-Adenosyl-L-methionine| + |Precorrin 4| <=> |S-Adenosyl-homocysteine| + |H+| + |Precorrin 5|	2.1.1.133	R05181	0
cobalt-precorrin-4 methyltransferase	|S-Adenosyl-L-methionine| + |Cobalt-precorrin 4| <=> |S-Adenosyl-homocysteine| + |Cobalt-precorrin	2.1.1.133	R05810	0
rxn06979	(2) |H2O| + (2) |ATP| + (2) |L-Glutamine| + |Cobyrinate| <=> (2) |ADP| + (2) |Phosphate| + (2) |L-Gluta	6.3.1.-	R05815	0
Hydrogenobyrinate <=> Hydrogenobyrinate a,c diamide	|H2O| + (2) |ATP| + (2) |L-Glutamine| + |Hydrogenobyrinate| <=> (2) |ADP| + |PPi| + (2) |L-Glutamate	Undetermined	None	0
adenosylcobyric acid synthase (glutamine-hydrolysing)	(4) |H2O| + (4) |ATP| + (4) |L-Glutamine| + |Adenosyl cobyrinate diamide| <=> (4) |ADP| + (4) |Phosph	6.3.5.10	R05225	0
Nicotinate-nucleotide:dimethylbenzimidazole	|Nicotinate ribonucleotide| + |Dimethylbenzimidazole| <=> |H+| + |Niacin| + |alpha-Ribazole 5'-p	2.4.2.21	R04148	1.25981
ATP:cob(I)yrinic acid-a,c-diamide Cobeta-adenosyltransferase	|ATP| + |Cob(I)yrinate diamide| <=> |Triphosphate| + |Adenosyl cobyrinate diamide|	2.5.1.17	R05220	0
cob(I)alamin adenosyltransferase	|ATP| + |Cbl| <=> |Calomide| + |Triphosphate|	2.5.1.17	R01492	0
ATP:cobinamide Cobeta-adenosyltransferase	|ATP| + |Cobinamide| <=> |Triphosphate| + |Adenosyl cobinamide|	2.5.1.17	R07268	1.25981
Succinyl-CoA:L-homoserine O-succinyltransferase	|Succinyl-CoA| + |L-Homoserine| => |CoA| + |O-Succinyl-L-homoserine|	2.3.1.46	R01777	0
Formate:tetrahydrofolate ligase (ADP-forming)	|ATP| + |Formate| + |Tetrahydrofolate| <=> |ADP| + |Phosphate| + |10-Formyltetrahydrofolate|	3.5.4.9,6.3.4.3	R00943	53.2107
5,10-Methenyltetrahydrofolate 5-hydrolase(decyclizing)	|H2O| + |5-10-Methenyltetrahydrofolate| <=> |H+| + |10-Formyltetrahydrofolate|	3.5.4.9,6.3.4.3	R01655	-51.9508
5,10-Methylenetetrahydrofolate:dUMP C-methyltransferase	|5-10-Methylenetetrahydrofolate| + |dUMP| => |dTMP| + |Dihydrofolate|	2.1.1.45	R02101	2.13204
5,6,7,8-Tetrahydrofolate:NADP+ oxidoreductase	|NADP| + |Tetrahydrofolate| <=> |NADPH| + |H+| + |Dihydrofolate|	1.5.1.3	R00939	-5.91147
Dihydrofolate:NADP+ oxidoreductase	|NADP| + |Dihydrofolate| <=> |NADPH| + |H+| + |Folate|	1.5.1.3	R02236	0
Adenosine 5'-monophosphate phosphohydrolase	|H2O| + |AMP| => |Phosphate| + |H+| + |Adenosine|	3.1.3.5,3.1.3.5)	R00183	0
Cytidine-5'-monophosphate phosphohydrolase	|H2O| + |CMP| => |Phosphate| + |H+| + |Cytidine|	3.1.3.5,3.1.3.5)	R00511	0
Uridine 5'-monophosphate phosphohydrolase	|H2O| + |UMP| => |Phosphate| + |H+| + |Uridine|	3.1.3.5,3.1.3.5)	R00963	0
Inosine 5'-monophosphate phosphohydrolase	|H2O| + |IMP| => |Phosphate| + |H+| + |Inosine|	3.1.3.5,3.1.3.5)	R01126	0
Guanosine 5'-monophosphate phosphohydrolase	|H2O| + |GMP| => |Phosphate| + |H+| + |Guanosine|	3.1.3.5,3.1.3.5)	R01227	0
Thymidylate 5'-phosphohydrolase	|H2O| + |dTMP| => |Phosphate| + |H+| + |Thymidine|	3.1.3.35,3.1.3.5,3.1.3.5)	R01569	0
2'-Deoxycytidine 5'-monophosphate phosphohydrolase	|H2O| + |dCMP| => |Phosphate| + |H+| + |Deoxycytidine|	3.1.3.5,3.1.3.5)	R01664	0
2'-Deoxyguanosine 5'-monophosphate phosphohydrolase	|H2O| + |dGMP| => |Phosphate| + |H+| + |Deoxyguanosine|	3.1.3.5,3.1.3.5)	R01968	0
2'-Deoxyadenosine 5'-monophosphate phosphohydrolase	|H2O| + |dAMP| => |Phosphate| + |H+| + |Deoxyadenosine|	3.1.3.5,3.1.3.5)	R02088	0
5'-nucleotidase (dUMP)	|H2O| + |dUMP| => |Phosphate| + |H+| + |Deoxyuridine|	3.1.3.5,3.1.3.5)	R02102	0
Nicotinamide ribonucleotide phosphohydrolase	|H2O| + |Nicotinamide ribonucleotide| => |Phosphate| + |H+| + |N-Ribosylnicotinamide|	3.1.3.5	R02323	0
Xanthosine 5'-phosphate phosphohydrolase	|H2O| + |XMP| => |Phosphate| + |H+| + |Xanthosine|	3.1.3.5,3.1.3.5)	R02719	0
Nicotinate D-ribonucleotide phosphohydrolase	|H2O| + |Nicotinate ribonucleotide| => |Phosphate| + |H+| + |Nicotinate D-ribonucleoside|	3.1.3.5	R03346	0
riboflavin transport in/out via proton symport	|H+[e]| + |Riboflavin[e]| <=> |H+| + |Riboflavin|	Undetermined	None	2.51962
hydrogenase (ubiquinone-8: 2 protons)	(2) |H+| + |H2| + |Ubiquinone-8| => (2) |H+[e]| + |Ubiquinol-8|	1.18.99.1	None	0
Hydrogenase (Demethylmenaquinone-8: 2 protons)	(2) |H+| + |H2| + |2-Demethylmenaquinone 8| => (2) |H+[e]| + |2-Demethylmenaquinol 8|	1.18.99.1	None	0
Hydrogenase (menaquinone8: 2 protons)	(2) |H+| + |H2| + |Menaquinone 8| => (2) |H+[e]| + |Menaquinol 8|	1.18.99.1	None	0
rxn03394	|NADPH| + |O2| + |H+| + |2-Octaprenylphenol| => |H2O| + |NADP| + |2-Octaprenyl-6-hydroxyphenol	Undetermined	R04987	1.25981
L-Arabinose:NAD+ 1-oxidoreductase	|NAD| + |L-Arabinose| <=> |NADH| + |H+| + |L-Arabinono-1,4-lactone|	1.1.1.46	R01757	0
GTP:alpha-D-mannose-1-phosphate guanylyltransferase	|GTP| + |D-Mannose1-phosphate| <=> |PPi| + |GDP-mannose|	2.7.7.13	R00885	0
D-Mannose 6-phosphate 1,6-phosphomutase	|D-mannose-6-phosphate| <=> |D-Mannose1-phosphate|	5.4.2.8	R01818	0
D-Glucosamine 1-phosphate 1,6-phosphomutase	|D-Glucosamine1-phosphate| <=> |D-Glucosamine phosphate|	5.4.2.10,5.4.2.2	R02060	-31.523
heptosyltransferase IV (LPS core synthesis)	|ADP-L-glycero-D-manno-heptose| + |glucosyl-glucosyl-galactosyl-glucosyl-inner core oligosaccharide lip	2.4.1.56	None	0
alpha,alpha-Trehalose-6-phosphate phosphoglucohydrolase	|H2O| + |Trehalose 6-phosphate| <=> |D-Glucose| + |D-glucose-6-phosphate|	3.2.1.122,3.2.1.93	R00837	-1000
nitric oxide, NADPH2:oxygen oxidoreductase	|NADPH| + (2) |O2| + (2) |NO| <=> |NADP| + |H+| + (2) |Nitrate|	1.14.12.17	R05725	-1000
nitric oxide, NAD(P)H2:oxygen oxidoreductase	|NADH| + (2) |O2| + (2) |NO| <=> |NAD| + |H+| + (2) |Nitrate|	1.14.12.17	R05724	1000
L-methionine:oxidized-thioredoxin S-oxidoreductase	|H2O| + |L-Methionine| + |trdox| <=> |L-Methionine S-oxide| + |trdrd|	1.8.4.13,1.8.4.14	R02025	0
10-Formyltetrahydrofolate amidohydrolase	|H2O| + |10-Formyltetrahydrofolate| => |Formate| + |H+| + |Tetrahydrofolate|	3.5.1.10	R00944	0
L-Glutamate 1-carboxy-lyase	|L-Glutamate| + |H+| => |CO2| + |GABA|	4.1.1.15,4.1.1.19	R00261	0
N-Carbamoyl-beta-alanine amidohydrolase	|H2O| + (2) |H+| + |3-Ureidopropanoate| => |CO2| + |NH3| + |beta-Alanine|	3.5.1.6	R00905	0
3-Ureidoisobutyrate amidohydrolase	|H2O| + (2) |H+| + |3-Ureidoisobutyrate| => |CO2| + |NH3| + |3-Aminoisobutanoate|	3.5.1.6	R04666	0
acetaldehyde:NAD+ oxidoreductase (CoA-acetylating)	|NAD| + |CoA| + |Acetaldehyde| <=> |NADH| + |Acetyl-CoA| + |H+|	1.2.1.10	R00228	339.988
methylmalonate-semialdehyde dehydrogenase (propanol)	|NAD| + |CoA| + |Propanal| <=> |NADH| + |H+| + |Propionyl-CoA|	1.2.1.27	None	0
Propane-1,2-diol hydro-lyase	|1,2-Propanediol| => |H2O| + |Propanal|	4.2.1.28	R02376	0
molybdate transport via ABC system	|H2O| + |ATP| + |Molybdate[e]| => |ADP| + |Phosphate| + |H+| + |Molybdate|	TC-3.A.1.8,3.A.1.8	None	0
ZN2t4	|Zn2+| + |H+[e]| + |K+[e]| <=> |Zn2+[e]| + |H+| + |K+|	Undetermined	None	-1.25981
cadminum transport out via antiport	|H+[e]| + |K+[e]| + |Cd2+| <=> |H+| + |K+| + |Cd2+[e]|	Undetermined	None	0
Maltotriose transport via ABC system	|H2O| + |ATP| + |Amylotriose[e]| => |ADP| + |Phosphate| + |H+| + |Amylotriose|	Undetermined	None	0
maltohexaose transport via ABC system	|H2O| + |ATP| + |Maltohexaose[e]| => |ADP| + |Phosphate| + |H+| + |Maltohexaose|	Undetermined	None	0
L-Lysine 2,3-aminomutase	|L-Lysine| <=> |L-beta-Lysine|	5.4.3.2	R00461	0
Carbonic acid hydro-lyase	|H+| + |H2CO3| <=> |H2O| + |CO2|	4.2.1.1	R00132	650.863
FMNH2:NADP+ oxidoreductase	|NADP| + |FMNH2| <= |NADPH| + |FMN| + |H+|	1.5.1.29	R05706	-1.25981
cadmium transport out via ABC system	|H2O| + |ATP| + |Cd2+| => |ADP| + |Phosphate| + |H+| + |Cd2+[e]|	TC-3.A.3,3.A.3	None	0
Mercury (Hg+2) ABC transporter	|H2O| + |ATP| + |Hg2+| => |ADP| + |Phosphate| + |H+| + |Hg2+[e]|	Undetermined	None	0
Lead (Pb+2) ABC transporter	|H2O| + |ATP| + |Pb| => |ADP| + |Phosphate| + |H+| + |Pb[e]|	Undetermined	None	0
Copper export via ATPase	|H2O| + |ATP| + |Cu2+| => |ADP| + |Phosphate| + |Cu2+[e]| + |H+|	TC-3.A.3,3.A.3	None	0
Copper transport via ABC system	|H2O| + |ATP| + |Cu2+[e]| => |ADP| + |Phosphate| + |Cu2+| + |H+|	Undetermined	None	1.25981
3-Methylcrotonoyl-CoA:carbon-dioxide ligase (ADP-forming)	|ATP| + |H2CO3| + |Dimethylacryloyl-CoA| <=> |ADP| + |Phosphate| + |H+| + |3-Methylglutaconyl-Co	6.4.1.4	R04138	0
(S)-3-Hydroxy-3-methylglutaryl-CoA hydro-lyase	|HMG-CoA| <=> |H2O| + |3-Methylglutaconyl-CoA|	4.2.1.18	R02085	0
(S)-3-Hydroxy-3-methylglutaryl-CoA acetoacetate-lyase	|HMG-CoA| <=> |Acetyl-CoA| + |Acetoacetate|	4.1.3.4	R01360	0
Butanoyl-CoA:oxygen 2-oxidoreductase	|FAD| + |Butyryl-CoA| <= |Crotonyl-CoA| + |FADH2|	1.3.3.6,1.3.99.2,1.3.99.3,	R01175	0
3-Methylbutanoyl-CoA:(acceptor) 2,3-oxidoreductase	|FAD| + |Isovaleryl-CoA| <= |FADH2| + |Dimethylacryloyl-CoA|	1.3.99.3,1.3.99.10	R04095	0
Deoxyadenosine:orthophosphate ribosyltransferase	|Phosphate| + |H+| + |Deoxyadenosine| <=> |Adenine| + |deoxyribose-1-phosphate|	2.4.2.1,2.4.2.4	R02557	-2.13204
inosine:orthophosphate ribosyltransferase	|Phosphate| + |H+| + |Inosine| <=> |HYXN| + |Ribose 1-phosphate|	2.4.2.1,2.4.2.15	R01863	0
Deoxyguanosine:orthophosphate ribosyltransferase	|Phosphate| + |H+| + |Deoxyguanosine| <=> |Guanine| + |deoxyribose-1-phosphate|	2.4.2.1,2.4.2.4	R01969	0
guanosine:orthophosphate ribosyltransferase	|Phosphate| + |H+| + |Guanosine| <=> |Guanine| + |Ribose 1-phosphate|	2.4.2.1,2.4.2.15	R02147	0
N-Ribosylnicotinamide:orthophosphate ribosyltransferase	|Phosphate| + |N-Ribosylnicotinamide| <=> |Nicotinamide| + |Ribose 1-phosphate|	2.4.2.1	R02294	0
Nicotinate D-ribonucleoside:orthophosphate ribosyltransferase	|Phosphate| + |Nicotinate D-ribonucleoside| <=> |Niacin| + |Ribose 1-phosphate|	2.4.2.1	R02295	0
Xanthosine:orthophosphate ribosyltransferase	|Phosphate| + |H+| + |Xanthosine| <=> |XAN| + |Ribose 1-phosphate|	2.4.2.1	R02297	0
Deoxyinosine:orthophosphate ribosyltransferase	|Phosphate| + |H+| + |Deoxyinosine| <=> |HYXN| + |deoxyribose-1-phosphate|	2.4.2.1,2.4.2.4	R02748	0
Adenosine:orthophosphate ribosyltransferase	|Phosphate| + |H+| + |Adenosine| <=> |Adenine| + |Ribose 1-phosphate|	2.4.2.1	R01561	7.55886
deoxyuridine:orthophosphate 2-deoxy-D-ribosyltransferase	|Phosphate| + |H+| + |Deoxyuridine| <=> |Uracil| + |deoxyribose-1-phosphate|	2.4.2.1,2.4.2.4,2.4.2.23,2	R02484	342.12
TRDR	|NADPH| + |H+| + |trdox| <=> |NADP| + |trdrd|	1.6.4.5,1.8.1.9	R02016	745.7
Biotin synthase	|S-Adenosyl-L-methionine| + |S| + |Dethiobiotin| => |L-Methionine| + |H+| + |BIOT| + |5'-Deoxy	2.8.1.6	None	0
magnesium transport in/out via permease (no H+)	|Mg| <=> |Mg[e]|	TC-1.A.35,1.A.35	None	-1.25981
cobalt transport in/out via permease (no H+)	|Co2+| <=> |Co2+[e]|	Undetermined	None	-1.25981
Xanthine ion-coupled transport	|H+[e]| + |XAN[e]| <=> |H+| + |XAN|	TC-2.A.40,2.A.40	None	1000
XMP:pyrophosphate phosphoribosyltransferase	|PPi| + |XMP| <= |PRPP| + |XAN|	2.4.2.22,2.4.2.8	R02142	0
(S)-Malate:NADP+ oxidoreductase(oxaloacetate-decarboxylating)	|NADP| + |L-Malate| => |NADPH| + |CO2| + |Pyruvate|	1.1.1.38,1.1.1.39,1.1.1.4	R00216	0
L-Malate glyoxylate-lyase (CoA-acetylating)	|CoA| + |H+| + |L-Malate| <= |H2O| + |Acetyl-CoA| + |Glyoxalate|	2.3.3.9,4.1.3.2	R00472	-3.77943
succinate transporter in/out via proton symport	|Succinate[e]| + |H+[e]| <=> |Succinate| + |H+|	TC-2.A.56,2.A.56	None	-1000
fumarate transport in/out via proton symport	|H+[e]| + |Fumarate[e]| <=> |H+| + |Fumarate|	TC-2.A.56,2.A.56	None	-724.485
(R)-Glycerate:NADP+ oxidoreductase	|NADP| + |Glycerate| <=> |NADPH| + |H+| + |Tartronate semialdehyde|	1.1.1.60	R01747	0
(R)-Glycerate:NAD+ oxidoreductase	|NAD| + |Glycerate| <=> |NADH| + |H+| + |Tartronate semialdehyde|	1.1.1.60	R01745	1000
Hydroxypyruvate ketol-isomerase	|Hydroxypyruvate| <=> |Tartronate semialdehyde|	5.3.1.22	R01394	-1000
O3-Acetyl-L-serine acetate-lyase (adding hydrogen sulfide)	|H2S| + |O-Acetyl-L-serine| => |Acetate| + |H+| + |L-Cysteine|	2.5.1.47,2.5.1.65,4.2.99.8	R00897	0
cysteine synthase (Thiosulfate)	|H2S2O3| + |O-Acetyl-L-serine| + |trdrd| => |Acetate| + |Sulfite| + |L-Cysteine| + |trdox|	4.2.99.8,2.5.1.47,2.5.1.49	R04859	0
succinate transport via proton symport (2 H)	|Succinate[e]| + (2) |H+[e]| <=> |Succinate| + (2) |H+|	Undetermined	None	0
Malate transport via proton symport (2 H)	(2) |H+[e]| + |L-Malate[e]| <=> (2) |H+| + |L-Malate|	Undetermined	None	664.289
Aspartate transport via proton symport (2 H)	|L-Aspartate[e]| + (2) |H+[e]| <=> |L-Aspartate| + (2) |H+|	Undetermined	None	1000
Fumarate transport via proton symport (2 H)	(2) |H+[e]| + |Fumarate[e]| <=> (2) |H+| + |Fumarate|	Undetermined	None	1000
L-lysine reversible transport via proton symport	|L-Lysine[e]| + |H+[e]| <=> |L-Lysine| + |H+|	Undetermined	None	51.3618
ethanolamine transport in/out via proton symport	|H+[e]| + |Aminoethanol[e]| <=> |H+| + |Aminoethanol|	TC-2.A.3.5,2.A.3.5	None	0
Ethanolamine ammonia-lyase	|Aminoethanol| => |NH3| + |Acetaldehyde|	4.3.1.7	R00749	0
3-Phospho-D-glycerate:NAD+ 2-oxidoreductase	|NAD| + |3-Phosphoglycerate| <=> |NADH| + |H+| + |3-Phosphonooxypyruvate|	1.1.1.95	R01513	0
L-Arabitol:NADP+ 1-oxidoreductase	|NADP| + |L-Lyxitol| <=> |NADPH| + |H+| + |L-Arabinose|	1.1.1.21	R01759	0
ATP:4-amino-5-hydroxymethyl-2-methylpyrimidine 5-phosphotransferase	|ATP| + |Toxopyrimidine| <=> |ADP| + |4-Amino-5-phosphomethyl-2-methylpyrimidine|	2.7.1.49	R03471	0
ATP:4-amino-2-methyl-5-phosphomethylpyrimidine phosphotransferase	|ATP| + |4-Amino-5-phosphomethyl-2-methylpyrimidine| <=> |ADP| + |4-Amino-2-methyl-5-diphospho	2.7.4.7	R04509	0
ATP:4-methyl-5-(2-hydroxyethyl)-thiazole 2-phosphotransferase	|ATP| + |4-Methyl-5--2-hydroxyethyl-thiazole| <=> |ADP| + |4-Methyl-5--2-phosphoethyl-thiazole|	2.7.1.50	R04448	0
4-hydroxy-2-oxopentanoate pyruvate-lyase	|Pyruvate| + |Acetaldehyde| <=> |4-Hydroxy-2-oxovalerate|	4.1.3.39	R00750	0
Urea-1-carboxylate amidohydrolase	|H2O| + (3) |H+| + |Allophanate| => (2) |CO2| + (2) |NH3|	3.5.1.54	R00005	0
iron (II) transport via ABC system	|H2O| + |ATP| + |Fe2+[e]| => |ADP| + |Phosphate| + |H+| + |Fe2+|	Undetermined	None	0
2-Phospho-D-glycerate 2,3-phosphomutase	|2-Phospho-D-glycerate| <=> |3-Phosphoglycerate|	5.4.2.1,5.4.2.4	R01518	-1000
Ammonia transport via diffusion	|NH3[e]| <=> |NH3|	Undetermined	None	-1000
IMP:NAD+ oxidoreductase	|H2O| + |NAD| + |IMP| <=> |NADH| + |H+| + |XMP|	1.1.1.205	R01130	12.6396
rxn07430	|TPP| + |H+| + |3-Methyl-2-oxobutanoate| => |CO2| + |2-Methyl-1-hydroxypropyl-TPP|	1.2.4.4	R07599	0
rxn07431	|Lipoamide| + |2-Methyl-1-hydroxypropyl-TPP| <=> |TPP| + |S-(2-Methylpropionyl)-dihydrolipoamide|	1.2.4.4	R07600	0
rxn07432	|TPP| + |H+| + |4MOP| => |CO2| + |3-Methyl-1-hydroxybutyl-TPP|	1.2.4.4	R07601	13.193
rxn07433	|Lipoamide| + |3-Methyl-1-hydroxybutyl-TPP| <=> |TPP| + |S-(3-Methylbutanoyl)-dihydrolipoamide-E|	1.2.4.4	R07602	13.193
rxn07434	|TPP| + |H+| + |3MOP| => |CO2| + |2-Methyl-1-hydroxybutyl-TPP|	1.2.4.4	R07603	13.193
rxn07435	|Lipoamide| + |2-Methyl-1-hydroxybutyl-TPP| <=> |TPP| + |S-(2-Methylbutanoyl)-dihydrolipoamide-E|	1.2.4.4	R07604	13.193
Acetyl-CoA:acetyl-CoA C-acetyltransferase	(2) |Acetyl-CoA| <=> |CoA| + |Acetoacetyl-CoA|	2.3.1.16,2.3.1.9	R00238	0
Propanoyl-CoA:acetyl-CoA C-acyltransferase	|Acetyl-CoA| + |Propionyl-CoA| <=> |CoA| + |2-Methylacetoacetyl-CoA|	2.3.1.16	R00927	0

Reactions with corresponding equation, EC number and FBA flux value under aerobic conditions in rich media

Acetyl-CoA:acetyl-CoA C-acetyltransferase	|Acetyl-CoA| + |Butyryl-CoA| <=> |CoA| + |3-Oxohexanoyl-CoA|	2.3.1.9,2.3.1.16	R01177	0
Octanoyl-CoA:acetyl-CoA C-acyltransferase	|Acetyl-CoA| + |Octanoyl-CoA| <=> |CoA| + |3-Oxodecanoyl-CoA|	2.3.1.16	R03778	0
myristoyl-CoA:acetylCoA C-myristoyltransferase	|Acetyl-CoA| + |Myristoyl-CoA| <=> |CoA| + |3-Oxopalmitoyl-CoA|	2.3.1.16,2.3.1.155	R03991	0
Decanoyl-CoA:acetyl-CoA C-acyltransferase	|Acetyl-CoA| + |Decanoyl-CoA| <=> |CoA| + |3-Oxododecanoyl-CoA|	2.3.1.16	R04742	0
Hexanoyl-CoA:acetyl-CoA C-acyltransferase	|Acetyl-CoA| + |Hexanoyl-CoA| <=> |CoA| + |3-Oxooctanoyl-CoA|	2.3.1.16	R04747	0
Lauroyl-CoA:acetyl-CoA C-acyltransferase	|Acetyl-CoA| + |Lauroyl-CoA| <=> |CoA| + |3-Oxotetradecanoyl-CoA|	2.3.1.16	R03858	0
Phenylacetaldehyde:NAD+ oxidoreductase	|H2O| + |NAD| + |alpha-Tolualdehyde| <=> |NADH| + |H+| + |PACT|	1.2.1.39,1.2.1.5	R02536	0
(S)-3-Hydroxybutanoyl-CoA hydro-lyase	|(S)-3-Hydroxybutyryl-CoA| <=> |H2O| + |Crotonyl-CoA|	4.2.1.17	R03026	0
(S)-3-Hydroxydodecanoyl-CoA hydro-lyase	|(S)-3-Hydroxydodecanoyl-CoA| <=> |H2O| + |(2E)-Dodecenoyl-CoA|	4.2.1.17,4.2.1.74	R04170	0
(S)-3-Hydroxyisobutyryl-CoA hydro-lyase	|H2O| + |Methacrylyl-CoA| <=> |(S)-3-Hydroxyisobutyryl-CoA|	4.2.1.17	R04224	0
(S)-3-Hydroxyhexadecanoyl-CoA hydro-lyase	|(S)-3-Hydroxyhexadecanoyl-CoA| <=> |H2O| + |(2E)-Hexadecenoyl-CoA|	4.2.1.17,4.2.1.74	R04738	0
(S)-3-Hydroxytetradecanoyl-CoA hydro-lyase	|(S)-3-Hydroxytetradecanoyl-CoA| <=> |H2O| + |(2E)-Tetradecenoyl-CoA|	4.2.1.17,4.2.1.74	R04740	0
(S)-Hydroxydecanoyl-CoA hydro-lyase	|(S)-Hydroxydecanoyl-CoA| <=> |H2O| + |(2E)-Decenoyl-CoA|	4.2.1.17,4.2.1.74	R04744	0
(S)-Hydroxyoctanoyl-CoA hydro-lyase	|(S)-Hydroxyoctanoyl-CoA| <=> |H2O| + |(2E)-Octenoyl-CoA|	4.2.1.17,4.2.1.74	R04746	0
(S)-Hydroxyhexanoyl-CoA hydro-lyase	|(S)-Hydroxyhexanoyl-CoA| <=> |H2O| + |(2E)-Hexenoyl-CoA|	4.2.1.17,4.2.1.74	R04749	0
(R)-2,3-Dihydroxy-3-methylpentanoate hydro-lyase	|2,3-Dihydroxy-3-methylvalerate| => |H2O| + |3MOP|	4.2.1.9	R05070	0
2,3-Dihydroxy-3-methylbutanoate hydro-lyase	|2,3-Dihydroxy-isovalerate| => |H2O| + |3-Methyl-2-oxobutanoate|	4.2.1.9	R01209,R04441	27.6337
ATP:L-glutamate 5-phosphotransferase	|ATP| + |L-Glutamate| + |H+| <=> |ADP| + |L-Glutamyl 5-phosphate|	2.7.2.11	R00239	0
L-Glutamate-5-semialdehyde:NADP+ 5-oxidoreductase (phosphorylationg)	|NADP| + |Phosphate| + |L-Glutamate5-semialdehyde| <= |NADPH| + |L-Glutamyl 5-phosphate|	1.2.1.41	R03313	0
ATP:shikimate 3-phosphotransferase	|ATP| + |Shikimate| <=> |ADP| + |3-phosphoshikimate|	2.7.1.71	R02412	7.55886
cytochrome oxidase bd (menaquinol-8: 2 protons)	(0.5) |O2| + (2) |H+| + |Menaquinol 8| => |H2O| + (2) |H+[e]| + |Menaquinone 8|	Undetermined	None	0
L-Glutamine amidohydrolase	|H2O| + |L-Glutamine| => |NH3| + |L-Glutamate|	3.5.1.2,3.5.1.38,6.3.5.5	R00256	0
6-Carboxyhexanoyl-CoA:L-alanine C-carboxyhexanoyltransferase	|L-Alanine| + |H+| + |Pimeloyl-CoA| => |CoA| + |CO2| + |8-Amino-7-oxononanoate|	2.3.1.47	R03210	0
L-Tryptophan:oxygen 2-oxidoreductase(decarboxylating)	|O2| + |L-Tryptophan| => |H2O| + |CO2| + |Indole-3-acetamide|	1.13.12.3	R00679	0
UTP:alpha-D-hexose-1-phosphate uridylyltransferase	|UTP| + |D-Galactose 1-phosphate| <=> |PPi| + |UDP-galactose|	2.7.7.10	R00502	0
UDPglucose 4-epimerase	|UDP-glucose| <=> |UDP-galactose|	5.1.3.2	R00291	0
UDP-N-acetyl-D-glucosamine 4-epimerase	|UDP-N-acetylglucosamine| <=> |UDP-N-acetyl-D-galactosamine|	5.1.3.7	R00418	8.59719
ATP:D-galactose 1-phosphotransferase	|ATP| + |Galactose| <=> |ADP| + |D-Galactose 1-phosphate|	2.7.1.6	R01092	0
alpha-D-Glucoside glucohydrolase	|H2O| + |Melibiose| <=> |D-Glucose| + |Galactose|	3.2.1.20,3.2.1.22	R01101	0
Raffinose galactohydrolase	|H2O| + |Melitose| => |Sucrose| + |Galactose|	3.2.1.22	R01103	0
a-galactosidase (stachyose)	|H2O| + |Stachyose| => |Galactose| + |Melitose|	3.2.1.22	R03634	0
rxn03838	|H2O| + |Manninotriose| => |Galactose| + |Melibiose|	3.2.1.22	R05549	0
glycerol-3-phosphate dehydrogenase (ubiquinone-8)	|Glycerol-3-phosphate| + |Ubiquinone-8| => |Glycerone-phosphate| + |Ubiquinol-8|	1.1.99.5	None	0
glycerol-3-phosphate dehydrogenase (menaquinone-8)	|Glycerol-3-phosphate| + |Menaquinone 8| => |Glycerone-phosphate| + |Menaquinol 8|	1.1.99.5	None	0
glycerol-3-phosphate dehydrogenase (demethylmenaquinone-8)	|Glycerol-3-phosphate| + |2-Demethylmenaquinone 8| => |Glycerone-phosphate| + |2-Demethylmena	1.1.99.5	None	0
sn-Glycerol-3-phosphate:(acceptor) 2-oxidoreductase	|FAD| + |Glycerol-3-phosphate| <=> |Glycerone-phosphate| + |FADH2|	1.1.99.5	R00848	1000
Peptidoglycan subunit synthesis	|Undecaprenyl-diphospho-N-acetylmuramoyl--N-acetylglucosamine-L-ala-D-glu-meso-2-6-diaminopimel	Undetermined	None	0.286573
L-Aspartate 1-carboxy-lyase	|L-Aspartate| + |H+| => |CO2| + |beta-Alanine|	4.1.1.11,4.1.1.15	R00489	2.51962
(R)-Pantoate:beta-alanine ligase (AMP-forming)	|ATP| + |beta-Alanine| + |Pantoate| => |PPi| + |AMP| + |PAN|	6.3.2.1	R02473	2.51962
5,10-Methylenetetrahydrofolate:3-methyl-2-oxobutanoate	|H2O| + |3-Methyl-2-oxobutanoate| + |5-10-Methylenetetrahydrofolate| <=> |Tetrahydrofolate| + |2-	2.1.2.11	R01226	2.51962
Biotin:CoA ligase (AMP-forming)	|ATP| + |H+| + |BIOT| <= |PPi| + |Biotinyl-5-AMP|	6.2.1.11,6.3.4.9,6.3.4.10,	R01074	0
RNA transcription	=> |RNA transcription|	Undetermined	None	158.673
Glycerone-phosphate phospho-lyase	|Glycerone-phosphate| => |Phosphate| + |H+| + |2-Oxopropanal|	4.2.3.3	R01016	0
2,3,4,5-Tetrahydrodipicolinate:NADP+ oxidoreductase	|NADP| + |tetrahydrodipicolinate| <=> |NADPH| + |H+| + |Dihydrodipicolinate|	1.3.1.26	R04199	-0.286573
cytochrome-c reductase (menaquinol 7: 3 protons)	|H+| + (2) |Cytochrome c3+| + |mql7| <=> (3) |H+[e]| + (2) |Cytochrome c2+| + |Menaquinone 7|	1.10.2.2,,1.10.2.2,	None	0
Phosphoenolpyruvate:3-phosphoshikimate	|Phosphoenolpyruvate| + |3-phosphoshikimate| <=> |Phosphate| + |H+| + |5-O--1-Carboxyvinyl-3-pho	2.5.1.19	R03460	7.55886
Prephenate:NAD+ oxidoreductase(decarboxylating)	|NAD| + |Prephenate| => |NADH| + |CO2| + |p-hydroxyphenylpyruvate|	1.3.1.12,1.3.1.43,1.3.1.5	R01728	0
Prephenate:NADP+ oxidoreductase(decarboxylating)	|NADP| + |Prephenate| => |NADPH| + |CO2| + |p-hydroxyphenylpyruvate|	1.3.1.13,1.3.1.43	R01730	0
L-Phenylalanine:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |L-Phenylalanine| <=> |L-Glutamate| + |Phenylpyruvate|	2.6.1.1,2.6.1.1 or 2.6.1.9,	R00694	0
L-Tyrosine:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |L-Tyrosine| <=> |L-Glutamate| + |p-hydroxyphenylpyruvate|	2.6.1.1,2.6.1.1 or 2.6.1.9,	R00734	0
L-Aspartate:2-oxoglutarate aminotransferase	|Oxaloacetate| + |Pretyrosine| <=> |L-Aspartate| + |Prephenate|	2.6.1.57,2.6.1.78	R01731	0
5-Amino-2-oxopentanoate:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |L-histidinol-phosphate| <=> |L-Glutamate| + |imidazole acetol-phosphate|	2.6.1.9	R03243	0
L-Serine hydro-lyase (adding indoleglycerol-phosphate)	|L-Serine| + |indol| => |H2O| + |L-Tryptophan|	4.2.1.20	R00674	0
(1S,2R)-1-C-(indol-3-yl)glycerol 3-phosphate	|Indoleglycerol phosphate| <=> |Glyceraldehyde3-phosphate| + |indol|	4.1.2.8,4.2.1.20	R02340	0
L-Serine hydro-lyase (adding indoleglycerol-phosphate)	|L-Serine| + |Indoleglycerol phosphate| => |H2O| + |L-Tryptophan| + |Glyceraldehyde3-phosphate|	4.2.1.20	R02722	0
N-(5-Phospho-beta-D-ribosyl)anthranilate ketol-isomerase	|N-5-phosphoribosyl-anthranilate| <=> |1-(2-carboxyphenylamino)-1-deoxyribulose 5-phosphate|	5.3.1.24	R03509	0
1-(2-Carboxyphenylamino)-1-deoxy-D-ribulose-5-phosphate	|H+| + |1-(2-carboxyphenylamino)-1-deoxyribulose 5-phosphate| => |H2O| + |CO2| + |Indoleglycerol p	4.1.1.48	R03508	0
N-(5-Phospho-D-ribosyl)anthranilate:pyrophosphate	|PPi| + |N-5-phosphoribosyl-anthranilate| <= |Anthranilate| + |PRPP|	2.4.2.18	R01073	0
Chorismate pyruvatemutase	|Chorismate| => |Prephenate|	5.4.99.5	R01715	0
2-Dehydro-3-deoxy-D-arabino-heptonate 7-phosphate phosphate-lyase	|DAHP| => |Phosphate| + |H+| + |5-Dehydroquinate|	4.2.3.4,4.6.1.3	R03083	7.55886
5-O-(1-Carboxyvinyl)-3-phosphoshikimate phosphate-lyase	|5-O--1-Carboxyvinyl-3-phosphoshikimate| => |Phosphate| + |H+| + |Chorismate|	4.2.3.5	R01714	7.55886
ATP:nucleoside-diphosphate phosphatransferase	|ATP| + |GDP| <=> |ADP| + |GTP|	2.7.4.6	R00330	-1000
ATP:nucleoside-diphosphate phosphatransferase	|ATP| + |IDP| <=> |ADP| + |ITP|	2.7.4.6	R00722	-1000
ATP:nucleoside-diphosphate phosphatransferase	|ATP| + |dADP| <=> |ADP| + |dATP|	2.7.4.6	R01137	-1000
ATP:nucleoside-diphosphate phosphatransferase	|ATP| + |dTDP| <=> |ADP| + |TTP|	2.7.4.6	R02093	-1000
ATP:nucleoside-diphosphate phosphatransferase	|ATP| + |dUDP| <=> |ADP| + |dUTP|	2.7.4.6	R02331	-1000
ATP:nucleoside-diphosphate phosphatransferase	|ATP| + |dGDP| <=> |ADP| + |dGTP|	2.7.4.6	R01857	-385.415
ATP:nucleoside-diphosphate phosphatransferase	|ATP| + |dCDP| <=> |ADP| + |dCTP|	2.7.4.6	R02326	2.13204
ATP:nucleoside-diphosphate phosphatransferase	|ATP| + |CDP| <=> |ADP| + |CTP|	2.7.4.6	R00570	997.868
ATP:nucleoside-diphosphate phosphatransferase	|ATP| + |UDP| <=> |ADP| + |UTP|	2.7.4.6	R00156	1000
trans,trans-Farnesyl-diphosphate:isopentenyl-diphosphate	|Isopentenyldiphosphate| + |Farnesyldiphosphate| => |PPi| + |Geranylgeranyl diphosphate|	2.5.1.29	R02061	3.77943
trans-hexaprenyltranstransferase	|Isopentenyldiphosphate| + |all-trans-Hexaprenyl diphosphate| => |PPi| + |all-trans-Heptaprenyl diphc	2.5.1.30	R05612	3.77943
PPTT	|Isopentenyldiphosphate| + |pendp| => |PPi| + |all-trans-Hexaprenyl diphosphate|	2.5.1.33	R05613	3.77943
GGTT	|Isopentenyldiphosphate| + |Geranylgeranyl diphosphate| => |PPi| + |pendp|	Undetermined	R07475	3.77943
rxn04674	|S-Adenosyl-L-methionine| + |Demethylphylloquinone| => |S-Adenosyl-homocysteine| + |H+| + |Vitam	2.1.1.-	R06859	0
UDP-L-rhamnose:flavonol-3-O-D-glucoside L-rhamnosyltransferase	|S-Adenosyl-L-methionine| + |2-Octaprenyl-6-methoxy-1,4-benzoquinone| => |S-Adenosyl-homocystei	2.1.1.-	R04990	1.25981
S-adenosylmethione:2-demethylmenaquinone methyltransferase	|S-Adenosyl-L-methionine| + |2-Demethylmenaquinone 8| => |S-Adenosyl-homocysteine| + |H+| + |M	Undetermined	None	1.25981
GTP 7,8-8,9-dihydrolase	|H2O| + |GTP| => |Formamidopyrimidine nucleoside triphosphate|	3.5.4.16	R00428	0
2-Amino-4-hydroxy-6-(erythro-1,2,3-trihydroxypropyl)	|H2O| + |7,8-Dihydroneopterin 3'-triphosphate| <= |2,5-Diamino-6-(5'-triphosphoryl-3&#39	3.5.4.16	R04639	0
Formamidopyrimidine nucleoside triphosphate 7,8-8,9-dihydrolase	|H2O| + |Formamidopyrimidine nucleoside triphosphate| <=> |Formate| + |H+| + |2,5-Diaminopyrimic	3.5.4.16	R05046	0
GTP 7,8-8,9-dihydrolase	|2,5-Diaminopyrimidine nucleoside triphosphate| <=> |2,5-Diamino-6-(5'-triphosphoryl-3',4	3.5.4.16,6.-.-.-	R05048	0
GTP 7,8-8,9-dihydrolase	|H2O| + |GTP| => |Formate| + |H+| + |7,8-Dihydroneopterin 3'-triphosphate|	3.5.4.16	R00424	3.77943
sn-Glycerol-3-phosphate:NAD+ 2-oxidoreductase	|NAD| + |Glycerol-3-phosphate| <=> |NADH| + |H+| + |Glycerone-phosphate|	1.1.1.8,1.1.1.94,1.1.1.26	R00842,R05679	-555.585
sn-Glycerol-3-phosphate:NADP+ 2-oxidoreductase	|NADP| + |Glycerol-3-phosphate| <=> |NADPH| + |H+| + |Glycerone-phosphate|	1.1.1.261,1.1.1.94	R05680,R00844	1000
1-tetradecanoyl-sn-glycerol 3-phosphate O-acyltransferase (n-C12:0)	|Dodecanoyl-ACP| + |1-dodecanoyl-sn-glycerol 3-phosphate| => |ACP| + |1,2-didodecanoyl-sn-glycerol	2.3.1.51	None	0
1-tetradecanoyl-sn-glycerol 3-phosphate O-acyltransferase (n-C14:0)	|Myristoyl-ACP| + |1-tetradecanoyl-sn-glycerol 3-phosphate| => |ACP| + |1,2-ditetradecanoyl-sn-glycer	2.3.1.51	None	0
1-tetradec-7-enoyl-sn-glycerol 3-phosphate O-acyltransferase (n-C14:1)	|Tetradecenoyl-ACP| + |1-tetradec-7-enoyl-sn-glycerol 3-phosphate| => |ACP| + |1,2-ditetradec-7-eno	2.3.1.51	None	0
1-hexadecanoyl-sn-glycerol 3-phosphate O-acyltransferase (n-C16:0)	|Palmitoyl-ACP| + |1-hexadecanoyl-sn-glycerol 3-phosphate| => |ACP| + |1,2-dihexadecanoyl-sn-glycer	2.3.1.51	None	0
1-hexadec-7-enoyl-sn-glycerol 3-phosphate O-acyltransferase (n-C16:1)	|Hexadecenoyl-ACP| + |1-hexadec-9-enoyl-sn-glycerol 3-phosphate| => |ACP| + |1,2-dihexadec-9-enoy	2.3.1.51	None	0
1-octadecanoyl-sn-glycerol 3-phosphate O-acyltransferase (n-C18:0)	|Octadecanoyl-ACP| + |1-octadecanoyl-sn-glycerol 3-phosphate| => |ACP| + |1,2-dioctadecanoyl-sn-gly	2.3.1.51	None	0
1-octadec-7-enoyl-sn-glycerol 3-phosphate O-acyltransferase (n-C18:1)	|Octadecenoyl-ACP| + |1-octadec-11-enoyl-sn-glycerol 3-phosphate| => |ACP| + |1,2-dioctadec-11-eno	2.3.1.51	None	0
palmitoyl-1-acylglycerol-3-phosphate O-acyltransferase	|Palmitoyl-CoA| + |1-hexadecanoyl-sn-glycerol 3-phosphate| => |CoA| + |1,2-dihexadecanoyl-sn-glycer	2.3.1.51	None	0
myristoyl-1-acylglycerol-3-phosphate O-acyltransferase	|Myristoyl-CoA| + |1-tetradecanoyl-sn-glycerol 3-phosphate| => |CoA| + |1,2-ditetradecanoyl-sn-glycer	2.3.1.51	None	0
isotetradecanoyl-1-acylglycerol-3-phosphate O-acyltransferase	|fa1coa| + |1-isotetradecanoyl-sn-glycerol 3-phosphate| => |CoA| + |1,2-diisotetradecanoyl-sn-glycero	2.3.1.51	None	0
isopentadecanoyl-1-acylglycerol-3-phosphate O-acyltransferase	|fa3coa| + |1-isopentadecanoyl-sn-glycerol 3-phosphate| => |CoA| + |1,2-diisopentadecanoyl-sn-glycer	2.3.1.51	None	0
anteisopentadecanoyl-1-acylglycerol-3-phosphate O-acyltransferase	|fa4coa| + |1-anteisopentadecanoyl-sn-glycerol 3-phosphate| => |CoA| + |1,2-dianteisopentadecanoyl	2.3.1.51	None	0
isohexadecanoyl-1-acylglycerol-3-phosphate O-acyltransferase	|fa6coa| + |1-isohexadecanoyl-sn-glycerol 3-phosphate| => |CoA| + |1,2-diisohexadecanoyl-sn-glycerol	2.3.1.51	None	0
stearoyl-1-acylglycerol-3-phosphate O-acyltransferase	|strcoa| + |1-octadecanoyl-sn-glycerol 3-phosphate| => |CoA| + |1,2-dioctadecanoyl-sn-glycerol 3-pho	2.3.1.51	None	6.59648
isoheptadecanoyl-1-acylglycerol-3-phosphate O-acyltransferase	|fa11coa| + |1-isoheptadecanoyl-sn-glycerol 3-phosphate| => |CoA| + |1,2-diisoheptadecanoyl-sn-glyc	2.3.1.51	None	6.59648
anteisoheptadecanoyl-1-acylglycerol-3-phosphate O-acyltransferase	|fa12coa| + |1-anteisoheptadecanoyl-sn-glycerol 3-phosphate| => |CoA| + |1,2-dianteisoheptadecanoy	2.3.1.51	None	6.59648
ITP:cytidine 5'-phosphotransferase	|ITP| + |Cytidine| <=> |CMP| + |IDP|	2.7.1.48	R00962	-1000
ATP:CMP phosphotransferase	|ATP| + |dCMP| <=> |ADP| + |dCDP|	2.7.4.14	R01665	0
ATP:CMP phosphotransferase	|ATP| + |CMP| <=> |ADP| + |CDP|	2.7.4.14	R00512	1000
L-Asparagine amidohydrolase	|H2O| + |L-Asparagine| => |NH3| + |L-Aspartate|	3.5.1.1,3.5.1.38	R00485	0
L-Glutamate:NAD+ oxidoreductase (deaminating)	|H2O| + |NAD| + |L-Glutamate| <=> |NADH| + |NH3| + |2-Oxoglutarate| + |H+|	1.4.1.2,1.4.1.3,1.4.1.4	R00243	-1000
L-Histidinol-phosphate phosphohydrolase	|H2O| + |L-histidinol-phosphate| => |Phosphate| + |H+| + |L-Histidinol|	3.1.3.15	R03013	0
Adenosylcobalamin 5'-phosphate synthase	|alpha-Ribazole| + |Adenosylcobinamide-GDP| <=> |GMP| + |Calomide|	2.7.8.26	R05223	1.25981
rxn04413	|GTP| + |Adenosyl cobinamide| <=> |GDP| + |H+| + |Adenosyl cobinamide phosphate|	2.7.1.156	R06558	-998.74
Adenosyl cobinamide phosphate guanyltransferase	|GTP| + |H+| + |Adenosyl cobinamide phosphate| <=> |PPi| + |Adenosylcobinamide-GDP|	2.7.7.62	R05222	1.25981
Adenosyl cobinamide kinase	|ATP| + |Adenosyl cobinamide| <=> |ADP| + |H+| + |Adenosyl cobinamide phosphate|	2.7.1.156	R05221	1000
L-threonine-O-3-phosphate carboxy-lyase	|H+| + |L-Threonine phosphate| => |CO2| + |(R)-1-Aminopropan-2-yl phosphate|	4.1.1.81	R06530	0
adenosylcobyric acid:(R)-1-aminopropan-2-ol ligase (ADP-forming)	|ATP| + |1-Aminopropan-2-ol| + |Adenosylcobyric acid| <=> |ADP| + |Phosphate| + |H+| + |Adenosyl	6.3.1.10	R07302	-1000
adenosylcobyric acid:(R)-1-aminopropan-2-yl phosphate ligase	|ATP| + |(R)-1-Aminopropan-2-yl phosphate| + |Adenosylcobyric acid| <=> |ADP| + |Phosphate| + (2)	6.3.1.10	R06529	0
adenosylcobinamide amidohydrolase	|1-Aminopropan-2-ol| + |Adenosylcobyric acid| <=> |H2O| + |Adenosyl cobinamide|	3.5.1.90	R05226	1000
Vitamin ABC transport	|H2O| + |ATP| + |Vitamin B12[e]| => |ADP| + |Phosphate| + |H+| + |Vitamin B12|	Undetermined	None	0
Cobalamin uptake in via ABC transport	|H2O| + |ATP| + |Vitamin B12r[e]| => |ADP| + |Phosphate| + |H+| + |Vitamin B12r|	Undetermined	None	0
L-Aspartate-4-semialdehyde hydro-lyase (adding pyruvate and	|Pyruvate| + |L-Aspartate4-semialdehyde| => (2) |H2O| + |H+| + |Dihydrodipicolinate|	4.2.1.52	R02292	0.286573
D-Glucosamine-6-phosphate ketol-isomerase(deaminating)	|H2O| + |D-Glucosamine phosphate| => |NH3| + |D-fructose-6-phosphate|	3.5.99.6	R00765	0

N-Acetyl-D-glucosamine-6-phosphate amidohydrolase	|H2O| + |N-Acetyl-D-glucosamine 6-phosphate| <=> |Acetate| + |D-Glucosamine phosphate|	3.5.1.25	R02059	-1000
6,7-Dimethyl-8-(1-D-ribityl)lumazine:6,7-dimethyl-8-(1-D-ribityl)	|4--1-D-Ribitylamino-5-aminouracil| + |3-4-dihydroxy-2-butanone4-phosphate| <=> (2) |H2O| + |Phosp	2.5.1.9	R04457	0
GTP 7,8-8,9-dihydrolase (pyrophosphate-forming)	(3) |H2O| + |GTP| => |PPi| + |Formate| + |H+| + |2,5-Diamino-6-(5'-phosphoribosylamino)-4-pyri	3.5.4.25	R00425	0
3,4-Dihydroxy-2-butanone 4-phosphate synthase	|D-Ribulose5-phosphate| => |Formate| + |H+| + |3-4-dihydroxy-2-butanone4-phosphate|	Undetermined	R07281	0
6,7-Dimethyl-8-(1-D-ribityl)lumazine:6,7-dimethyl-8-(1-D-ribityl)	(2) |6-7-Dimethyl-8--1-D-ribityllumazine| => |Riboflavin| + |4--1-D-Ribitylamino-5-aminouracil|	2.5.1.9	R00066	0
5-amino-6-(5-phosphoribitylamino)uracil:NADP+ 1'-oxidoreductase	|NADP| + |5-Amino-6--5-phosphoribitylaminouracil| <=> |NADPH| + |H+| + |5-Amino-6--5-phosphorib	1.1.1.193	R03458	0
2,5-Diamino-6-hydroxy-4-(5-phosphoribosylamino)-pyrimidine	|H2O| + |H+| + |2,5-Diamino-6-(5'-phosphoribosylamino)-4-pyrimidineone| => |NH3| + |5-Amino	3.5.4.26	R03459	0
meso-2,6-Diaminoheptanedioate carboxy-lyase	|H+| + |meso-2,6-Diaminopimelate| => |CO2| + |L-Lysine|	4.1.1.20	R00451	0
Cytidine:orthophosphate alpha-D-ribosyltransferase	|Phosphate| + |H+| + |Cytidine| <=> |Cytosine| + |Ribose 1-phosphate|	2.4.2.2	R02296	-829.041
Uridine:orthophosphate ribosyltransferase	|Phosphate| + |H+| + |Uridine| <=> |Uracil| + |Ribose 1-phosphate|	2.4.2.2,2.4.2.3	R01876	474.033
D-Ribose 1,5-phosphomutase	|Ribose 1-phosphate| <=> |ribose-5-phosphate|	5.4.2.2,5.4.2.7	R01057	-347.449
2-Deoxy-D-ribose 1-phosphate 1,5-phosphomutase	|deoxyribose-1-phosphate| <=> |deoxyribose-5-phosphate|	5.4.2.7	R02749	339.988
ADPribose ribophosphohydrolase	|H2O| + |ADPribose| => |AMP| + |ribose-5-phosphate|	3.6.1.13	None	0
L-Proline:NAD+ 5-oxidoreductase	|NAD| + |L-Proline| <=> |NADH| + |H+| + |1-Pyrroline-5-carboxylate|	1.5.1.2	R01248	-1000
L-Proline:NADP+ 5-oxidoreductase	|NADP| + |L-Proline| <=> |NADPH| + |H+| + |1-Pyrroline-5-carboxylate|	1.5.1.2	R01251	1000
beta-D-Glucose-6-phosphate:NADP+ 1-oxoreductase	|NADP| + |beta-D-Glucose 6-phosphate| <=> |NADPH| + |H+| + |6-phospho-D-glucono-1-5-lactone|	1.1.1.49	R02736	-1000
D-Glucose-6-phosphate:NADP+ 1-oxoreductase	|NADP| + |D-glucose-6-phosphate| <=> |NADPH| + |H+| + |6-phospho-D-glucono-1-5-lactone|	1.1.1.49	R00835	1000
Acetyl-CoA:carbon-dioxide ligase (ADP-forming)	|Acetyl-CoA| + |Carboxybiotin-carboxyl-carrier protein| => |Malonyl-CoA| + |Holo-[carboxylase]|	6.4.1.2	R04386	263.859
(R)-2-Methyl-3-oxopropanoyl-CoA 2-epimerase	|L-methylmalonyl-CoA| <=> |D-methylmalonyl-CoA|	5.1.99.1	R02765	0
(S)-2-methylbutanoyl-CoA:enzyme N6-(dihydrolipoyl)lysine	|Dihydrolipoamide| + |2-Methylbutyryl-CoA| <=> |CoA| + |S-(2-Methylbutanoyl)-dihydrolipoamide-E|	2.3.1.168	R03174	-13.193
3-methylbutanoyl-CoA:enzyme N6-(dihydrolipoyl)lysine	|Dihydrolipoamide| + |Isovaleryl-CoA| <=> |CoA| + |S-(3-Methylbutanoyl)-dihydrolipoamide-E|	2.3.1.168	R04097	-13.193
2-methylpropanoyl-CoA:enzyme N6-(dihydrolipoyl)lysine	|Dihydrolipoamide| + |Isobutyryl-CoA| <=> |CoA| + |S-(2-Methylpropionyl)-dihydrolipoamide|	2.3.1.168	R02662	0
ATP:butyrate 1-phosphotransferase	|ATP| + |H+| + |Butyrate| <=> |ADP| + |Butanoylphosphate|	2.7.2.7	R01688	0
L-Valine:NAD+ oxidoreductase(deaminating)	|H2O| + |NAD| + |L-Valine| <=> |NADH| + |NH3| + |H+| + |3-Methyl-2-oxobutanoate|	1.4.1.9	R01434	-1000
L-leucine:NAD+ oxidoreductase(deaminating)	|H2O| + |NAD| + |L-Leucine| <=> |NADH| + |NH3| + |H+| + |4MOP|	1.4.1.9	R01088	204.157
L-Isoleucine:NAD+ oxidoreductase(deaminating)	|H2O| + |NAD| + |L-Isoleucine| <=> |NADH| + |NH3| + |H+| + |3MOP|	1.4.1.9	R02196	1000
Butanoyl-CoA:orthophosphate butanoyltransferase	|Phosphate| + |H+| + |Butyryl-CoA| <=> |CoA| + |Butanoylphosphate|	2.3.1.19	R01174	0
1-Deoxy-D-xylulose-5-phosphate pyruvate-lyase (carboxylating)	|Pyruvate| + |H+| + |Glyceraldehyde3-phosphate| => |CO2| + |1-deoxy-D-xylulose5-phosphate|	2.2.1.7	R05636	33.3877
5,10-methylenetetrahydrofolate:NADP+ oxidoreductase	|NADP| + |5-10-Methylenetetrahydrofolate| <=> |NADPH| + |5-10-Methenyltetrahydrofolate|	1.5.1.5	R01220	-51.9508
biotin-carboxyl-carrier-protein:carbon-dioxide ligase (ADP-forming)	|ATP| + |H2CO3| + |Holo-[carboxylase]| <=> |ADP| + |Phosphate| + |H+| + |Carboxybiotin-carboxyl-ca	6.3.4.14	R04385	263.859
3-Dehydroquinate hydro-lyase	|5-Dehydroquinate| => |H2O| + |3-Dehydroshikimate|	4.2.1.10	R03084	7.55886
glycine:lipoylprotein oxidoreductase (decarboxylating and	|Glycine| + |H+| + |Lipoylprotein| => |CO2| + |S-Aminomethyldihydrolipoylprotein|	1.4.4.2	R03425	0
ATP:D-glucose 6-phosphotransferase	|ATP| + |beta-D-Glucose| <=> |ADP| + |beta-D-Glucose 6-phosphate|	2.7.1.1,2.7.1.2	R01600	-1000
ATP:D-glucose 6-phosphotransferase	|ATP| + |D-Glucose| <=> |ADP| + |D-glucose-6-phosphate|	2.7.1.1,2.7.1.2	R00299,R01786	1000
2'-Deoxyuridine 5'-diphosphate:oxidized-thioredoxin	|H2O| + |dUDP| + |trdox| <= |UDP| + |trdrd|	1.17.4.1	R02018	-344.252
2'-Deoxycytidine diphosphate:oxidized-thioredoxin 2'-oxidoreduc	|H2O| + |dCDP| + |trdox| <= |CDP| + |trdrd|	1.17.4.1	R02024	-2.13204
Ribonucleotide reductase: ADP	|ADP| + |trdrd| => |H2O| + |dADP| + |trdox|	1.17.4.1	R02017	0
Ribonucleotide reductase: GDP	|GDP| + |trdrd| => |H2O| + |dGDP| + |trdox|	1.17.4.1	R02019	2.13204
5-Formyltetrahydrofolate cyclo-ligase (ADP-forming)	|ATP| + |5-Formyltetrahydrofolate| => |ADP| + |Phosphate| + |5-10-Methenyltetrahydrofolate|	6.3.3.2	R02301	0
Superoxide:superoxide oxidoreductase	(2) |O2-| => |O2| + |H2O2|	1.15.1.1	R00275	0
Orthophosphate-ABC transport	|H2O| + |ATP| + |Phosphate[e]| <=> |ADP| + (2) |Phosphate| + |H+|	TC-3.A.1.7,3.A.1.7	None	-1000
MECDPDH	|NADH| + |H+| + |2-C-methyl-D-erythritol2-4-cyclodiphosphate| => |H2O| + |NAD| + |1-Hydroxy-2-me	Undetermined	None	33.3877
Zinc-ABC transport	|H2O| + |ATP| + |Zn2+[e]| => |ADP| + |Phosphate| + |Zn2+| + |H+|	TC-3.A.1.15,3.A.1.15	None	0
(R)-Pantoate:NADP+ 2-oxidoreductase	|NADP| + |Pantoate| <=> |NADPH| + |H+| + |2-Dehydropantoate|	1.1.1.169	R02472	-2.51962
isopentenyl-diphosphate:NAD(P)+ oxidoreductase	|NADPH| + |H+| + |1-Hydroxy-2-methyl-2-butenyl 4-diphosphate| => |H2O| + |NADP| + |Isopentenyld	1.17.1.2	R05884	0
dimethyallyl diphosphate:NADP+ oxidoreductase	|H2O| + |NADP| + |DMAPP| <= |NADPH| + |H+| + |1-Hydroxy-2-methyl-2-butenyl 4-diphosphate|	1.17.1.2	R07219	0
1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase (dmpp)	|NADH| + |H+| + |1-Hydroxy-2-methyl-2-butenyl 4-diphosphate| => |H2O| + |NAD| + |DMAPP|	Undetermined	R08210	4.066
1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase (ipdp)	|NADH| + |H+| + |1-Hydroxy-2-methyl-2-butenyl 4-diphosphate| => |H2O| + |NAD| + |Isopentenyldipl	Undetermined	R08209	29.3217
Cytidine aminohydrolase	|H2O| + |H+| + |Cytidine| => |NH3| + |Uridine|	3.5.4.5	R01878	0
Deoxycytidine aminohydrolase	|H2O| + |H+| + |Deoxycytidine| => |NH3| + |Deoxyuridine|	3.5.4.14,3.5.4.5	R02485	0
diacylglycerol kinase (n-C18:0)	|ATP| + |1,2-Diacyl-sn-glycerol dioctadecanoyl| <=> |ADP| + |1,2-dioctadecanoyl-sn-glycerol 3-phospha	2.7.1.107	None	-1.14629
isoheptadecanoyl-Diacylglycerol kinase	|ATP| + |1,2-Diisoheptadecanoyl-sn-glycerol| <=> |ADP| + |1,2-diisoheptadecanoyl-sn-glycerol 3-phosp	2.7.1.107	None	-1.14629
anteisoheptadecanoyl-Diacylglycerol kinase	|ATP| + |1,2-Dianteisoheptadecanoyl-sn-glycerol| <=> |ADP| + |1,2-dianteisoheptadecanoyl-sn-glycero	2.7.1.107	None	-1.14629
diacylglycerol kinase (n-C12:0)	|ATP| + |1,2-Diacyl-sn-glycerol didodecanoyl| <=> |ADP| + |1,2-didodecanoyl-sn-glycerol 3-phosphate|	2.7.1.107	None	0
diacylglycerol kinase (n-C14:0)	|ATP| + |1,2-Diacyl-sn-glycerol ditetradecanoyl| <=> |ADP| + |1,2-ditetradecanoyl-sn-glycerol 3-phosph	2.7.1.107	None	0
diacylglycerol kinase (n-C14:1)	|ATP| + |1,2-Diacyl-sn-glycerol ditetradec-7-enoyl| <=> |ADP| + |1,2-ditetradec-7-enoyl-sn-glycerol 3-p	2.7.1.107	None	0
diacylglycerol kinase (n-C16:0)	|ATP| + |1,2-Diacyl-sn-glycerol dihexadecanoyl| <=> |ADP| + |1,2-dihexadecanoyl-sn-glycerol 3-phosph	2.7.1.107	None	0
diacylglycerol kinase (n-C16:1)	|ATP| + |1,2-Diacyl-sn-glycerol dihexadec-9-enoyl| <=> |ADP| + |1,2-dihexadec-9-enoyl-sn-glycerol 3-pl	2.7.1.107	None	0
diacylglycerol kinase (n-C18:1)	|ATP| + |1,2-Diacyl-sn-glycerol dioctadec-11-enoyl| <=> |ADP| + |1,2-dioctadec-11-enoyl-sn-glycerol 3-	2.7.1.107	None	0
isotetradecanoyl-Diacylglycerol kinase	|ATP| + |1,2-Diisotetradecanoyl-sn-glycerol| <=> |ADP| + |1,2-diisotetradecanoyl-sn-glycerol 3-phosph	2.7.1.107	None	0
isopentadecanoyl-Diacylglycerol kinase	|ATP| + |1,2-Diisopentadecanoyl-sn-glycerol| <=> |ADP| + |1,2-diisopentadecanoyl-sn-glycerol 3-phosp	2.7.1.107	None	0
anteisopentadecanoyl-Diacylglycerol kinase	|ATP| + |1,2-Dianteisopentadecanoyl-sn-glycerol| <=> |ADP| + |1,2-dianteisopentadecanoyl-sn-glycero	2.7.1.107	None	0
isohexadecanoyl-Diacylglycerol kinase	|ATP| + |1,2-Diisohexadecanoyl-sn-glycerol| <=> |ADP| + |1,2-diisohexadecanoyl-sn-glycerol 3-phospha	2.7.1.107	None	0
PIt8	|Phosphate[e]| + (3) |Na+[e]| <=> |Phosphate| + (3) |Na+|	TC-2.A.58,2.A.58	None	333.333
2-Deoxy-D-ribose-5-phosphate acetaldehyde-lyase	|deoxyribose-5-phosphate| <=> |Acetaldehyde| + |Glyceraldehyde3-phosphate|	4.1.2.4	R01066	339.988
dCMP aminohydrolase	|H2O| + |H+| + |dCMP| => |NH3| + |dUMP|	3.5.4.12	R01663	0
ATP:nicotinamide-nucleotide adenylyltransferase	|ATP| + |Nicotinate ribonucleotide| <=> |PPi| + |Deamido-NAD|	2.7.7.1,2.7.7.18	R03005	2.51962
Shikimate:NADP+ 5-oxidoreductase	|NADP| + |Shikimate| <=> |NADPH| + |H+| + |3-Dehydroshikimate|	1.1.1.25,1.1.1.282	R02413	-7.55886
Phosphatidylserine decarboxylase (n-C12:0)	(2) |H+| + |phosphatidylserine didodecanoyl| <=> |CO2| + |phosphatidylethanolamine didodecanoyl|	4.1.1.65	None	0
Phosphatidylserine decarboxylase (n-C14:0)	(2) |H+| + |phosphatidylserine ditetradecanoyl| <=> |CO2| + |phosphatidylethanolamine ditetradecanc	4.1.1.65	None	0
Phosphatidylserine decarboxylase (n-C14:1)	(2) |H+| + |phosphatidylserine ditetradec-7-enoyl| <=> |CO2| + |phosphatidylethanolamine ditetradec-	4.1.1.65	None	0
Phosphatidylserine decarboxylase (n-C16:0)	(2) |H+| + |phosphatidylserine dihexadecanoyl| <=> |CO2| + |phosphatidylethanolamine dihexadecanc	4.1.1.65	None	0
Phosphatidylserine decarboxylase (n-C16:1)	(2) |H+| + |phosphatidylserine dihexadec-9-enoyl| <=> |CO2| + |phosphatidylethanolamine dihexadec-	4.1.1.65	None	0
Phosphatidylserine decarboxylase (n-C18:1)	(2) |H+| + |phosphatidylserine dioctadec-11-enoyl| <=> |CO2| + |phosphatidylethanolamine dioctadec-	4.1.1.65	None	0
isotetradecanoyl-phosphatidylserine decarboxylase	(2) |H+| + |Diisotetradecanoylphosphatidylserine| <=> |CO2| + |Diisotetradecanoylphosphatidylethanc	4.1.1.65	None	0
isopentadecanoyl-phosphatidylserine decarboxylase	(2) |H+| + |Diisopentadecanoylphosphatidylserine| <=> |CO2| + |Diisopentadecanoylphosphatidylethai	4.1.1.65	None	0
anteisopentadecanoyl-phosphatidylserine decarboxylase	(2) |H+| + |Dianteisopentadecanoylphosphatidylserine| <=> |CO2| + |Dianteisopentadecanoylphospha	4.1.1.65	None	0
isohexadecanoyl-phosphatidylserine decarboxylase	(2) |H+| + |Diisohexadecanoylphosphatidylserine| <=> |CO2| + |Diisohexadecanoylphosphatidylethano	4.1.1.65	None	0
Phosphatidylserine decarboxylase (n-C18:0)	(2) |H+| + |phosphatidylserine dioctadecanoyl| <=> |CO2| + |phosphatidylethanolamine dioctadecanoy	4.1.1.65	None	1.36255
isoheptadecanoyl-phosphatidylserine decarboxylase	(2) |H+| + |Diisoheptadecanoylphosphatidylserine| <=> |CO2| + |Diisoheptadecanoylphosphatidylethai	4.1.1.65	None	1.36255
anteisoheptadecanoyl-phosphatidylserine decarboxylase	(2) |H+| + |Dianteisoheptadecanoylphosphatidylserine| <=> |CO2| + |Dianteisoheptadecanoylphospha	4.1.1.65	None	1.36255
Phosphatidylserine syntase (n-C12:0)	|L-Serine| + |CDP-1,2-didodecanoylglycerol| => |CMP| + |H+| + |phosphatidylserine didodecanoyl|	2.7.8.8	None	0
Phosphatidylserine syntase (n-C14:0)	|L-Serine| + |CDP-1,2-ditetradecanoylglycerol| => |CMP| + |H+| + |phosphatidylserine ditetradecanoyl	2.7.8.8	None	0
Phosphatidylserine syntase (n-C14:1)	|L-Serine| + |CDP-1,2-ditetradec-7-enoylglycerol| => |CMP| + |H+| + |phosphatidylserine ditetradec-7-	2.7.8.8	None	0
Phosphatidylserine syntase (n-C16:0)	|L-Serine| + |CDP-1,2-dihexadecanoylglycerol| => |CMP| + |H+| + |phosphatidylserine dihexadecanoyl	2.7.8.8	None	0
Phosphatidylserine syntase (n-C16:1)	|L-Serine| + |CDP-1,2-dihexadec-9-enoylglycerol| => |CMP| + |H+| + |phosphatidylserine dihexadec-9-	2.7.8.8	None	0
Phosphatidylserine syntase (n-C18:1)	|L-Serine| + |CDP-1,2-dioctadec-11-enoylglycerol| => |CMP| + |H+| + |phosphatidylserine dioctadec-1	2.7.8.8	None	0
isotetradecanoyl-CDPdiacylglycerol-serine O-phosphatidyltransferase	|L-Serine| + |CDP-1,2-diisotetradecanoylglycerol| => |CMP| + |H+| + |Diisotetradecanoylphosphatidyls	2.7.8.8	None	0
isopentadecanoyl-CDPdiacylglycerol-serine O-phosphatidyltransferase	|L-Serine| + |CDP-1,2-diisopentadecanoylglycerol| => |CMP| + |H+| + |Diisopentadecanoylphosphatidy	2.7.8.8	None	0
anteisopentadecanoyl-CDPdiacylglycerol-serine O-phosphatidyltransferase	|L-Serine| + |CDP-1,2-dianteisopentadecanoylglycerol| => |CMP| + |H+| + |Dianteisopentadecanoylph	2.7.8.8	None	0
isohexadecanoyl-CDPdiacylglycerol-serine O-phosphatidyltransferase	|L-Serine| + |CDP-1,2-diisohexadecanoylglycerol| => |CMP| + |H+| + |Diisohexadecanoylphosphatidyls	2.7.8.8	None	0
Phosphatidylserine syntase (n-C18:0)	|L-Serine| + |CDP-1,2-dioctadecanoylglycerol| => |CMP| + |H+| + |phosphatidylserine dioctadecanoyl|	2.7.8.8	None	1.36255
isoheptadecanoyl-CDPdiacylglycerol-serine O-phosphatidyltransferase	|L-Serine| + |CDP-1,2-diisoheptadecanoylglycerol| => |CMP| + |H+| + |Diisoheptadecanoylphosphatidy	2.7.8.8	None	1.36255
anteisoheptadecanoyl-CDPdiacylglycerol-serine O-phosphatidyltransferase	|L-Serine| + |CDP-1,2-dianteisoheptadecanoylglycerol| => |CMP| + |H+| + |Dianteisoheptadecanoylph	2.7.8.8	None	1.36255
Undecaprenyl-diphosphate phosphohydrolase	|H2O| + |Bactoprenyl diphosphate| => |Phosphate| + |H+| + |Undecaprenylphosphate|	3.6.1.27	R05627	0.286573
GLUt4i	|L-Glutamate[e]| + |Na+[e]| <=> |L-Glutamate| + |Na+|	Undetermined	None	-1000
GLUt2	|L-Glutamate[e]| + |H+[e]| <=> |L-Glutamate| + |H+|	Undetermined	None	141.518
L-Cystathionine Lysteine-lyase (deaminating)	|H2O| + |Cystathionine| => |NH3| + |L-Cysteine| + |2-Oxobutyrate|	4.4.1.1	R01001	0
L-Cysteine L-homocysteine-lyase (deaminating)	|H2O| + |L-Cysteine| => |NH3| + |Pyruvate| + |H2S|	4.1.99.1,4.4.1.1,4.4.1.8	R00782	0
L-Serine hydro-lyase (adding homocysteine)	|L-Serine| + |Homocysteine| <=> |H2O| + |Cystathionine|	4.2.1.22	R01290,R01289	0
S-Adenosyl-L-homocysteine homocysteinylribohydrolase	|H2O| + |S-Adenosyl-homocysteine| <=> |Adenine| + |S-Ribosylhomocysteine|	3.2.2.9	R00194	0
Methylthioadenosine methylthioribohydrolase	|H2O| + |5-Methylthioadenosine| <=> |Adenine| + |5-Methylthio-D-ribose|	3.2.2.16,3.2.2.9	R01401	0
S-Adenosyl-L-methionine:DNA (cytosine-5-)-methyltransferase	|S-Adenosyl-L-methionine| + |DNA cytosine| => |S-Adenosyl-homocysteine| + |H+| + |DNA 5-methylcy	2.1.1.37	R04858	0
GTP:uridine 5'-phosphotransferase	|GTP| + |Uridine| <=> |GDP| + |UMP|	2.7.1.48	R00968	-1000
dATP:cytidine 5'-phosphotransferase	|dATP| + |Cytidine| <=> |CMP| + |dADP|	2.7.1.48	R01548	-1000
dTTP:cytidine 5'-phosphotransferase	|TTP| + |Cytidine| <=> |CMP| + |dTDP|	2.7.1.48	R02096	-1000
dTTP:uridine 5'-phosphotransferase	|Uridine| + |TTP| <=> |UMP| + |dTDP|	2.7.1.48	R02097	-1000
dCTP:uridine 5'-phosphotransferase	|Uridine| + |dCTP| <=> |UMP| + |dCDP|	2.7.1.48	R02327	-1000
dUTP:cytidine 5'-phosphotransferase	|dUTP| + |Cytidine| <=> |CMP| + |dUDP|	2.7.1.48	R02372	-1000
dGTP:uridine 5'-phosphotransferase	|dGTP| + |Uridine| <=> |UMP| + |dGDP|	2.7.1.48	R01880	-387.547
ITP:uridine 5'-phosphotransferase	|ITP| + |Uridine| <=> |IDP| + |UMP|	2.7.1.48	R00970	0
dUTP:uridine 5'-phosphotransferase	|Uridine| + |dUTP| <=> |UMP| + |dUDP|	2.7.1.48	R02332	0
UTP:cytidine 5'-phosphotransferase	|UTP| + |Cytidine| <=> |UDP| + |CMP|	2.7.1.48	R00516	857.217
UTP:uridine 5'-phosphotransferase	|UTP| + |Uridine| <=> |UDP| + |UMP|	2.7.1.48	R00967	929.546
GTP:cytidine 5'-phosphotransferase	|GTP| + |Cytidine| <=> |GDP| + |CMP|	2.7.1.48	R00517	971.824
dATP:uridine 5'-phosphotransferase	|dATP| + |Uridine| <=> |UMP| + |dADP|	2.7.1.48	R01549	983.969
ATP:cytidine 5'-phosphotransferase	|ATP| + |Cytidine| <=> |ADP| + |CMP|	2.7.1.48	R00513	1000
ATP:uridine 5'-phosphotransferase	|ATP| + |Uridine| <=> |ADP| + |UMP|	2.7.1.48	R00964	1000

Reactions with corresponding equation, EC number and FBA flux value under aerobic conditions in rich media

dGTP:cytidine 5'-phosphotransferase	|dGTP| + |Cytidine| <=> |CMP| + |dGDP|	2.7.1.48	R02091	1000
dCTP:cytidine 5'-phosphotransferase	|dCTP| + |Cytidine| <=> |CMP| + |dCDP|	2.7.1.48	R02371	1000
Thiamin-ABC transport	|H2O| + |ATP| + |Thiamin[e]| => |ADP| + |Phosphate| + |H+| + |Thiamin|	Undetermined	None	0
ATP:GTP 3'-pyrophosphotransferase	|ATP| + |GTP| <=> |AMP| + |Guanosine 5'-triphosphate,3'-diphosphate|	2.7.6.5	R00429	0
AMP:pyrophosphate phosphoribosyltransferase	|PPi| + |AMP| <= |PRPP| + |Adenine|	2.4.2.7,2.4.2.8	R00190	-5.42682
2-Methyl-4-amino-5-hydroxymethylpyrimidine-diphosphate:4-methyl-5-	|H+| + |4-Methyl-5--2-phosphoethyl-thiazole| + |4-Amino-2-methyl-5-diphosphomethylpyrimidine| <=	2.5.1.3	R03223	0
Aminoacetic acid:oxygen oxidoreductase (deaminating)	|H2O| + |O2| + |Glycine| => |NH3| + |H2O2| + |Glyoxalate|	1.4.3.3,1.4.3.19	R00366	0
glycine oxidase	|Glycine| => (3) |H+| + |Iminoglycine|	1.4.3.19	R07463	0
Deamino-NAD+:ammonia ligase (AMP-forming)	|ATP| + |NH3| + |Deamido-NAD| => |NAD| + |PPi| + |AMP|	6.3.1.5	R00189	2.51962
quinolinate synthase	(2) |H2O| + |Phosphate| + |H+| + |Quinolinate| <= |Glycerone-phosphate| + |Iminoaspartate|	Undetermined	R04292	-2.51962
Nicotinate-nucleotide:pyrophosphate phosphoribosyltransferase	|CO2| + |PPi| + |Nicotinate ribonucleotide| <= (2) |H+| + |PRPP| + |Quinolinate|	2.4.2.19	R03348	-2.51962
L-Aspartic acid:oxygen oxidoreductase (deaminating)	|H2O| + |O2| + |L-Aspartate| => |NH3| + |H2O2| + |Oxaloacetate|	1.4.3.2,1.4.3.16	R00357	0
L-aspartate oxidase	|O2| + |L-Aspartate| => |H2O2| + |H+| + |Iminoaspartate|	1.4.3.16	R00481	2.51962
Prephenate hydro-lyase (decarboxylating)	|H+| + |Prephenate| => |H2O| + |CO2| + |Phenylpyruvate|	4.2.1.51,4.2.1.91	R01373	0
thiazole phosphate synthesis	|ATP| + |L-Tyrosine| + |L-Cysteine| + |1-deoxy-D-xylulose5-phosphate| => |H2O| + |CO2| + |PPi| + |A	Undetermined	None	0
Tetrahydrofolate:L-glutamate gamma-ligase (ADP-forming)	|ATP| + |L-Glutamate| + |Tetrahydrofolate| => |ADP| + |Phosphate| + |H+| + |THF-L-glutamate|	6.3.2.17	R00942	0
10-Formyltetrahydrofolate:L-glutamate ligase (ADP-forming)	|ATP| + |L-Glutamate| + |10-Formyltetrahydrofolate| => |ADP| + |Phosphate| + |H+| + |10-Formyl-TH	6.3.2.12,6.3.2.17	R01654	0
7,8-dihydropteroate:L-glutamate ligase (ADP-forming)	|ATP| + |L-Glutamate| + |Dihydropteroate| => |ADP| + |Phosphate| + |H+| + |Dihydrofolate|	6.3.2.12,6.3.2.17	R02237	3.77943
(S)-4-Amino-5-oxopentanoate 4,5-aminomutase	|5-Aminolevulinate| <=> |L-Glutamate1-semialdehyde|	5.4.3.8	R02272	-20.157
5-Aminolevulinate hydro-lyase(adding 5-aminolevulinate and	(2) |5-Aminolevulinate| => (2) |H2O| + |H+| + |Porphobilinogen|	4.2.1.24	R00036	10.0785
Hydroxymethylbilane hydro-lyase(cyclizing)	|Hydroxymethylbilane| <=> |H2O| + |Uroporphyrinogen III|	4.2.1.75	R03165	2.51962
Porphobilinogen ammonia-lyase (polymerizing)	|H2O| + (4) |Porphobilinogen| => (4) |NH3| + |Hydroxymethylbilane|	2.5.1.61,4.3.1.8	R00084	2.51962
L-glutamate-semialdehyde: NADP+	|NADPH| + |H+| + |L-Glutamyl-tRNA-Glu| <=> |NADP| + |L-Glutamate1-semialdehyde| + |tRNA-Glu|	1.2.1.70	R04109	20.157
2-Isopropylmalate hydro-lyase	|2-Isopropylmalate| <=> |H2O| + |2-Isopropylmaleate|	4.2.1.33	R03968	0
3-Isopropylmalate hydro-lyase	|3-Isopropylmalate| <=> |H2O| + |2-Isopropylmaleate|	4.2.1.33	R04001	0
3-Isopropylmalate:NAD+ oxidoreductase	|NAD| + |3-Isopropylmalate| <=> |NADH| + |H+| + |2-isopropyl-3-oxosuccinate|	1.1.1.85	R04426	0
3-Carboxy-3-hydroxy-4-methylpentanoate 3-methyl-2-oxobutanoate-lyase	|CoA| + |H+| + |2-Isopropylmalate| <= |H2O| + |Acetyl-CoA| + |3-Methyl-2-oxobutanoate|	2.3.3.13,4.1.3.12	R01213	0
(R)-2,3-Dihydroxy-3-methylbutanoate:NADP+ oxidoreductase	|NADP| + |2,3-Dihydroxy-isovalerate| <=> |NADPH| + |H+| + |2-Oxo-3-hydroxyisovalerate|	1.1.1.86	R04440	-1000
2,3-Dihydroxy-3-methylbutanoate:NADP+ oxidoreductase (isomerizing)	|NADPH| + |H+| + |ALCTT| <=> |NADP| + |2,3-Dihydroxy-isovalerate|	1.1.1.86	R03051,R04439	-972.366
(R)-2,3-Dihydroxy-3-methylpentanoate:NADP+ oxidoreductase	|NADP| + |2,3-Dihydroxy-3-methylvalerate| <=> |NADPH| + |H+| + |(R)-3-Hydroxy-3-methyl-2-oxopent	1.1.1.86	R05068	0
(S)-2-Aceto-2-hydroxybutanoate:NADP+ oxidoreductase (isomerizing)	|2-Aceto-2-hydroxybutanoate| <=> |(R)-3-Hydroxy-3-methyl-2-oxopentanoate|	1.1.1.86,5.4.99.3	R05069	0
2-Acetolactate methylmutase	|ALCTT| <=> |2-Oxo-3-hydroxyisovalerate|	5.4.99.3,1.1.1.86	R05071,R03052	1000
2-Acetolactate pyruvate-lyase (carboxylating)	|TPP| + |ALCTT| <=> |Pyruvate| + |2-Hydroxyethyl-ThPP|	2.2.1.6	R03050,R04672	-27.6337
(S)-2-Aceto-2-hydroxybutanoate pyruvate-lyase (carboxylating)	|2-Oxobutyrate| + |2-Hydroxyethyl-ThPP| <=> |TPP| + |2-Aceto-2-hydroxybutanoate|	2.2.1.6	R04673	0
L-Isoleucine:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |L-Isoleucine| <=> |L-Glutamate| + |3MOP|	2.6.1.42	R02199	-986.807
L-Leucine:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |L-Leucine| <=> |L-Glutamate| + |4MOP|	2.6.1.42,2.6.1.6,2.6.1.67	R01090	-190.964
L-Valine:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |L-Valine| <=> |L-Glutamate| + |3-Methyl-2-oxobutanoate|	2.6.1.42,2.6.1.6	R01214	974.886
L-Glutamate racemase	|L-Glutamate| <=> |D-Glutamate|	5.1.1.3	R00260	1000
fumarate reductase	|Fumarate| + |2-Demethylmenaquinol 8| <=> |Succinate| + |2-Demethylmenaquinone 8|	1.3.99.1	None	2.51962
fumarate reductase	|Fumarate| + |Menaquinol 8| <=> |Succinate| + |Menaquinone 8|	1.3.99.1	None	997.48
succinate dehydrogenase (irreversible)	|Succinate| + |Ubiquinone-8| <=> |Fumarate| + |Ubiquinol-8|	1.3.99.1	None	1000
Succinate:(acceptor) oxidoreductase	|FAD| + |Succinate| <= |Fumarate| + |FADH2|	1.3.99.1	R00408	-1000
succinate dehyrdogenase	|FADH2| + |Ubiquinone-8| => |FAD| + |Ubiquinol-8|	Undetermined	None	0
hydrogen peroxide reductase (thioredoxin)	|H2O2| + |trdrd| => (2) |H2O| + |trdox|	Undetermined	None	398.443
L-Arabinose-ABC transport	|H2O| + |ATP| + |L-Arabinose[e]| => |ADP| + |Phosphate| + |H+| + |L-Arabinose|	3.A.1.2	None	0
Maltose-ABC transport	|H2O| + |ATP| + |Maltose[e]| => |ADP| + |Phosphate| + |H+| + |Maltose|	3.A.1.1	None	0
S-Adenosyl-L-methionine carboxy-lyase	|S-Adenosyl-L-methionine| + |H+| => |CO2| + |S-Adenosylmethioninamine|	4.1.1.50	R00178	0
D-Glyceraldehyde-3-phosphate:NADP+ oxidoreductase(phosphorylating)	|NADP| + |Phosphate| + |Glyceraldehyde3-phosphate| <= |NADPH| + |1,3-Bisphospho-D-glycerate|	1.2.1.59	R01063	20.1654
ATP:dephospho-CoA 3'-phosphotransferase	|ATP| + |Dephospho-CoA| => |ADP| + |CoA|	2.7.1.24	R00130	2.51962
(S)-malate:NAD+ oxidoreductase	|NAD| + |L-Malate| <=> |NADH| + |Oxaloacetate| + |H+|	1.1.1.37	R00342	924.343
Oxalosuccinate:NADP+ oxidoreductase (decarboxylating)	|H+| + |Oxalosuccinate| => |CO2| + |2-Oxoglutarate|	1.1.1.42	R00268	0
Isocitrate:NADP+ oxidoreductase (decarboxylating)	|NADP| + |Isocitrate| <=> |NADPH| + |H+| + |Oxalosuccinate|	1.1.1.42	R01899	0
Citrate oxaloacetate-lyase ((pro-3S)-CH2COO- -> acetyl-CoA)	|CoA| + |H+| + |Citrate| <= |H2O| + |Acetyl-CoA| + |Oxaloacetate|	2.3.3.1,2.3.3.3,4.1.3.7	R00351	0
ATP:pyruvate O2-phosphotransferase	|ATP| + |Pyruvate| <=> |ADP| + |Phosphoenolpyruvate|	2.7.1.40	R00200	-967.551
ATP:Sedoheptulose 7-phosphate 1-phosphotransferase	|ATP| + |Sedoheptulose7-phosphate| <=> |ADP| + |Sedoheptulose 1,7-bisphosphate|	2.7.1.11	R01843	-1000
ATP:D-fructose-6-phosphate 1-phosphotransferase	|ATP| + |D-fructose-6-phosphate| <=> |ADP| + |D-fructose-1,6-bisphosphate|	2.7.1.11	R00756,R04779	558.104
(3R)-3-Hydroxypalmitoyl-[acyl-carrier-protein]:NADP+ oxidoreductase	|NADP| + |R-3-hydroxypalmitoyl-acyl-carrierprotein-| <=> |NADPH| + |H+| + |3-oxohexadecanoyl-acp|	1.1.1.100,2.3.1.85,2.3.1.8	R04543	-13.193
(3R)-3-Hydroxyhexanoyl-[acyl-carrier-protein]:NADP+ oxidoreductase	|NADP| + |D-3-Hydroxyhexanoyl-[acp]| <=> |NADPH| + |H+| + |3-Oxohexanoyl-[acp]|	1.1.1.100,2.3.1.85,2.3.1.8	R04953	-13.193
(3R)-3-Hydroxydecanoyl-[acyl-carrier-protein]:NADP+ oxidoreductase	|NADP| + |(R)-3-Hydroxydecanoyl-[acyl-carrier protein]| <=> |NADPH| + |H+| + |3-oxodecanoyl-acp|	1.1.1.100,2.3.1.85,2.3.1.8	R04534	-13.193
(3R)-3-Hydroxybutanoyl-[acyl-carrier protein]:NADP+ oxidoreductase	|NADP| + |(R)-3-Hydroxybutanoyl-[acyl-carrier protein]| <=> |NADPH| + |H+| + |Acetoacetyl-ACP|	1.1.1.100,2.3.1.85,2.3.1.8	R04533	-13.193
(3R)-3-Hydroxydodecanoyl-[acyl-carrier-protein]:NADP+ oxidoreductase	|NADP| + |D-3-Hydroxydodecanoyl-[acp]| <=> |NADPH| + |H+| + |3-oxododecanoyl-acp|	1.1.1.100,2.3.1.85,2.3.1.8	R04964	-13.193
(3R)-3-Hydroxyoctanoyl-[acyl-carrier-protein]:NADP+ oxidoreductase	|NADP| + |(R)-3-Hydroxyoctanoyl-[acyl-carrier protein]| <=> |NADPH| + |H+| + |3-oxooctanoyl-acp|	1.1.1.100,2.3.1.85,2.3.1.8	R04536	-13.193
(3R)-3-Hydroxytetradecanoyl-[acyl-carrier-protein]:NADP+	|NADP| + |HMA| <=> |NADPH| + |H+| + |3-oxotetradecanoyl-acp|	1.1.1.100,2.3.1.85,2.3.1.8	R04566	-13.193
4-methyl-3-oxo-pentanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |4-methyl-3-oxo-pentanoyl-ACP| <=> |NADP| + |4-methyl-3-hydroxy-pentanoyl-ACP	1.1.1.0	None	0
6-methyl-3-oxo-heptanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |6-methyl-3-oxo-heptanoyl-ACP| <=> |NADP| + |6-methyl-3-hydroxy-heptanoyl-ACP	1.1.1.0	None	0
8-methyl-3-oxo-nonanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |8-methyl-3-oxo-nonanoyl-ACP| <=> |NADP| + |8-methyl-3-hydroxy-nonanoyl-ACP|	1.1.1.0	None	0
10-methyl-3-oxo-undecanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |10-methyl-3-oxo-undecanoyl-ACP| <=> |NADP| + |10-methyl-3-hydroxy-undecanoy	1.1.1.0	None	0
12-methyl-3-oxo-tridecanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |12-methyl-3-oxo-tridecanoyl-ACP| <=> |NADP| + |12-methyl-3-hydroxy-tridecanoy	1.1.1.0	None	0
14-methyl-3-oxo-pentadecanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |14-methyl-3-oxo-pentadecanoyl-ACP| <=> |NADP| + |14-methyl-3-hydroxy-pentad	1.1.1.0	None	0
4-methyl-3-oxo-hexanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |4-methyl-3-oxo-hexanoyl-ACP| <=> |NADP| + |4-methyl-3-hydroxy-hexanoyl-ACP|	1.1.1.0	None	13.193
6-methyl-3-oxo-octanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |6-methyl-3-oxo-octanoyl-ACP| <=> |NADP| + |6-methyl-3-hydroxy-octanoyl-ACP|	1.1.1.0	None	13.193
8-methyl-3-oxo-decanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |8-methyl-3-oxo-decanoyl-ACP| <=> |NADP| + |8-methyl-3-hydroxy-decanoyl-ACP|	1.1.1.0	None	13.193
10-methyl-3-oxo-dodecanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |10-methyl-3-oxo-dodecanoyl-ACP| <=> |NADP| + |10-methyl-3-hydroxy-dodecanoy	1.1.1.0	None	13.193
12-methyl-3-oxo-tetra-decanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |12-methyl-3-oxo-tetra-decanoyl-ACP| <=> |NADP| + |12-methyl-3-hydroxy-tetra-de	1.1.1.0	None	13.193
14-methyl-3-oxo-hexa-decanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |14-methyl-3-oxo-hexa-decanoyl-ACP| <=> |NADP| + |14-methyl-3-hydroxy-hexa-de	1.1.1.0	None	13.193
5-methyl-3-oxo-hexanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |5-methyl-3-oxo-hexanoyl-ACP| <=> |NADP| + |5-methyl-3-hydroxy-hexanoyl-ACP|	1.1.1.0	None	13.193
7-methyl-3-oxo-octanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |7-methyl-3-oxo-octanoyl-ACP| <=> |NADP| + |7-methyl-3-hydroxy-octanoyl-ACP|	1.1.1.0	None	13.193
9-methyl-3-oxo-decanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |9-methyl-3-oxo-decanoyl-ACP| <=> |NADP| + |9-methyl-3-hydroxy-decanoyl-ACP|	1.1.1.0	None	13.193
11-methyl-3-oxo-dodecanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |11-methyl-3-oxo-dodecanoyl-ACP| <=> |NADP| + |11-methyl-3-hydroxy-dodecanoy	1.1.1.0	None	13.193
13-methyl-3-oxo-tetra-decanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |13-methyl-3-oxo-tetra-decanoyl-ACP| <=> |NADP| + |13-methyl-3-hydroxy-tetra-de	1.1.1.0	None	13.193
15-methyl-3-oxo-hexa-decanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |15-methyl-3-oxo-hexa-decanoyl-ACP| <=> |NADP| + |15-methyl-3-hydroxy-hexa-de	1.1.1.0	None	13.193
3-Hydroxyoctodecanoyl-ACP:NADP+ oxidoreductase	|NADPH| + |H+| + |3-Oxooctodecanoyl-ACP| <=> |NADP| + |3-Hydroxyoctodecanoyl-ACP|	1.1.1.0	None	13.193
N-(L-Argininosuccinate) arginie-lyase	|L-Argininosuccinate| <=> |L-Arginine| + |Fumarate|	4.3.2.1	R01086	0
L-Citrulline:L-aspartate ligase (AMP-forming)	|ATP| + |L-Aspartate| + |Citrulline| <=> |PPi| + |AMP| + |L-Argininosuccinate|	6.3.4.5	R01954	0
ATP:acetate phosphotransferase	|ATP| + |H+| + |Propionate| <=> |ADP| + |Propionyl phosphate|	2.7.2.1,2.7.2.15	R01353	-1000
ATP:acetate phosphotransferase	|ATP| + |Acetate| + |H+| <=> |ADP| + |Acetylphosphate|	2.7.2.1,2.7.2.15	R00315	-865.74
ATP:NAD+ 2'-phosphotransferase	|ATP| + |NAD| <=> |NADP| + |ADP|	2.7.1.23	R00104	-996.608
NAD kinase (dTTP)	|NAD| + |TTP| <=> |NADP| + |dTDP|	2.7.1.23	None	997.868
Phosphoenolpyruvate:D-erythrose-4-phosphate	|H2O| + |Phosphoenolpyruvate| + |D-Erythrose4-phosphate| => |Phosphate| + |H+| + |DAHP|	2.5.1.54,4.1.2.15	R01826	7.55886
UDP-N-acetylmuramate:L-alanine ligase (ADP-forming)	|ATP| + |L-Alanine| + |UDP-MurNAc| => |ADP| + |Phosphate| + |H+| + |UDP-N-acetylmuramoyl-L-alar	6.3.2.8	R03193	0.286573
3-phenylpropionate transport via proton symport, reversible	|H+[e]| + |Phenylpropanoate[e]| <=> |H+| + |Phenylpropanoate|	Undetermined	None	0
Lysophospholipase L1 (2-acylglycerophosphotidate, n-C12:0) (periplasm)	|H2O| + |1-dodecanoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |ddca|	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphotidate, n-C14:0) (periplasm)	|H2O| + |1-tetradecanoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |Myristic acid	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphotidate, n-C14:1) (periplasm)	|H2O| + |1-tetradec-7-enoyl-sn-glycerol 3-phosphate| => |H+| + |Glycerol-3-phosphate| + |tetradecen	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphotidate, n-C16:0) (periplasm)	|H2O| + |1-hexadecanoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |Palmitate|	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphotidate, n-C16:1) (periplasm)	|H2O| + |1-hexadec-9-enoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |hexadecer	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphotidate, n-C18:0) (periplasm)	|H2O| + |1-octadecanoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |ocdca|	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphotidate, n-C18:1) (periplasm)	|H2O| + |1-octadec-11-enoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |octadece	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoethanolamine, n-C12:0) (periple	|H2O| + |1-Acyl-sn-glycero-3-phosphoethanolamine dodecanoyl| <=> |H+| + |Glycerophosphoethanola	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoethanolamine, n-C14:0) (periple	|H2O| + |1-Acyl-sn-glycero-3-phosphoethanolamine tetradecanoyl| <=> |H+| + |Glycerophosphethan	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoethanolamine, n-C14:1) (periple	|H2O| + |1-Acyl-sn-glycero-3-phosphoethanolamine tetradec-7-enoyl| <=> |H+| + |Glycerophosphoeth	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoethanolamine, n-C16:0) (periple	|H2O| + |1-Acyl-sn-glycero-3-phosphoethanolamine hexadecanoyl| <=> |H+| + |Palmitate| + |Glycero	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoethanolamine, n-C16:1) (periple	|H2O| + |1-Acyl-sn-glycero-3-phosphoethanolamine hexadec-9-enoyl| <=> |H+| + |Glycerophosphoeth	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoethanolamine, n-C18:0) (periple	|H2O| + |H+| + |L-2-Lysophosphatidylethanolamine| => |Glycerophosphoethanolamine| + |ocdca|	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoethanolamine, n-C18:1) (periple	|H2O| + |L-2-Lysophosphatidylethanolamine| <=> |H+| + |Glycerophosphoethanolamine| + |octadecer	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoglycerol, n-C12:0) (periplasm)	|H2O| + |1-Acyl-sn-glycero-3-phosphoglycerol dodecanoyl| <=> |H+| + |ddca| + |Glycerophosphoglyce	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoglycerol, n-C14:0) (periplasm)	|H2O| + |1-Acyl-sn-glycero-3-phosphoglycerol tetradecanoyl| <=> |H+| + |Glycerophosphoglycerol| + |	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoglycerol, n-C14:1) (periplasm)	|H2O| + |1-Acyl-sn-glycero-3-phosphoglycerol tetradec-7-enoyl| <=> |H+| + |Glycerophosphoglycerol|	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoglycerol, n-C16:0) (periplasm)	|H2O| + |1-Acyl-sn-glycero-3-phosphoglycerol hexadecanoyl| <=> |H+| + |Palmitate| + |Glycerophosph	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoglycerol, n-C16:1) (periplasm)	|H2O| + |1-Acyl-sn-glycero-3-phosphoglycerol hexadec-9-enoyl| <=> |H+| + |Glycerophosphoglycerol|	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoglycerol, n-C18:0) (periplasm)	|H2O| + |1-Acyl-sn-glycero-3-phosphoglycerol octadecanoyl| <=> |H+| + |ocdca| + |Glycerophosphogly	3.1.1.5	None	0
Lysophospholipase L1 (2-acylglycerophosphoglycerol, n-C18:1) (periplasm)	|H2O| + |1-Acyl-sn-glycero-3-phosphoglycerol octadec-11-enoyl| <=> |H+| + |Glycerophosphoglycerol|	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphotidate, n-C12:0)	|H2O| + |2-dodecanoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |ddca|	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphotidate, n-C14:0)	|H2O| + |2-tetradecanoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |Myristic acid	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphotidate, n-C14:1)	|H2O| + |2-tetradec-7-enoyl-sn-glycerol 3-phosphate| => |H+| + |Glycerol-3-phosphate| + |tetradecen	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphotidate, n-C16:0)	|H2O| + |2-hexadecanoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |Palmitate|	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphotidate, n-C16:1)	|H2O| + |2-hexadec-9-enoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |hexadecer	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphotidate, n-C18:0)	|H2O| + |2-octadecanoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |ocdca|	3.1.1.5	None	0

Lysophospholipase L2 (2-acylglycerophosphotidate, n-C18:1)	|H2O| + |2-octadec-11-enoyl-sn-glycerol 3-phosphate| <=> |H+| + |Glycerol-3-phosphate| + |octadece	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoethanolamine, n-C12:0)	|H2O| + |2-Acyl-sn-glycero-3-phosphoethanolamine dodecanoyl| <=> |H+| + |Glycerophosphoethanola	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoethanolamine, n-C14:0)	|H2O| + |2-Acyl-sn-glycero-3-phosphoethanolamine tetradecanoyl| <=> |H+| + |Glycerophosphoethan	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoethanolamine, n-C14:1)	|H2O| + |2-Acyl-sn-glycero-3-phosphoethanolamine tetradec-7-enoyl| => |H+| + |Glycerophosphoetha	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoethanolamine, n-C16:0)	|H2O| + |2-Acyl-sn-glycero-3-phosphoethanolamine (n-C16:0)| <=> |H+| + |Palmitate| + |Glycerophos	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoethanolamine, n-C16:1)	|H2O| + |2-Acyl-sn-glycero-3-phosphoethanolamine hexadec-9-enoyl| <=> |H+| + |Glycerophosphoeth	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoethanolamine, n-C18:0)	|H2O| + |2-Acyl-sn-glycero-3-phosphoethanolamine (n-C18:0)| <=> |H+| + |Glycerophosphoethanolam	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoethanolamine, n-C18:1)	|H2O| + |2-Acyl-sn-glycero-3-phosphoethanolamine octadec-11-enoyl| <=> |H+| + |Glycerophosphoeth	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoglycerol, n-C12:0)	|H2O| + |2-Acyl-sn-glycero-3-phosphoglycerol dodecanoyl| <=> |H+| + |ddca| + |Glycerophosphoglyce	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoglycerol, n-C14:0)	|H2O| + |2-Acyl-sn-glycero-3-phosphoglycerol tetradecanoyl| <=> |H+| + |Glycerophosphoglycerol| + |	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoglycerol, n-C14:1)	|H2O| + |2-Acyl-sn-glycero-3-phosphoglycerol tetradec-7-enoyl| => |H+| + |Glycerophosphoglycerol| +	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoglycerol, n-C16:0)	|H2O| + |2-Acyl-sn-glycero-3-phosphoglycerol hexadecanoyl| <=> |H+| + |Palmitate| + |Glycerophosph	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoglycerol, n-C16:1)	|H2O| + |2-Acyl-sn-glycero-3-phosphoglycerol hexadec-9-enoyl| <=> |H+| + |Glycerophosphoglycerol|	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoglycerol, n-C18:0)	|H2O| + |2-Acyl-sn-glycero-3-phosphoglycerol octadecanoyl| <=> |H+| + |ocdca| + |Glycerophosphogly	3.1.1.5	None	0
Lysophospholipase L2 (2-acylglycerophosphoglycerol, n-C18:1)	|H2O| + |2-Acyl-sn-glycero-3-phosphoglycerol octadec-11-enoyl| <=> |H+| + |Glycerophosphoglycerol|	3.1.1.5	None	0
ATP:L-methione S-adenosyltransferase	|H2O| + |ATP| + |L-Methionine| => |Phosphate| + |PPi| + |S-Adenosyl-L-methionine|	2.5.1.6	R00177	8.81867
ATP:oxaloacetate carboxy-lyase (transphosphorylating)	|ATP| + |Oxaloacetate| + |H+| => |ADP| + |CO2| + |Phosphoenolpyruvate|	4.1.1.49	R00341	0
Fe(II):oxygen oxidoreductase	|O2| + (4) |H+| + (4) |Fe2+| <=> (2) |H2O| + (4) |fe3|	1.16.3.1	R00078	-0.944858
S-(5-deoxy-D-ribos-5-yl)-L-homocysteine homocysteine-lyase	|S-Ribosylhomocysteine| => |Homocysteine| + |4-5-dihydroxy-2-3-pentanedione|	4.4.1.21	None	0
coproporphyrinogen-III:S-adenosyl-L-methionine	(2) |S-Adenosyl-L-methionine| + |CoproporphyrinogenIII| <=> (2) |CO2| + (2) |L-Methionine| + |Protop	1.3.99.22	R06895	0
O-Succinylbenzoate:CoA ligase (AMP-forming)	|ATP| + |CoA| + |H+| + |Succinylbenzoate| => |PPi| + |AMP| + |Succinylbenzoyl-CoA|	6.2.1.26	R04030	2.51962
rxn05024	|H+| + |Succinylbenzoyl-CoA| => |H2O| + |1,4-Dihydroxy-2-naphthoyl-CoA|	4.1.3.36	R07263	0
O-Succinylbenzoyl-CoA 1,4-dihydroxy-2-naphthoate-lyase	|Succinylbenzoyl-CoA| => |CoA| + |1-4-Dihydroxy-2-naphthoate|	4.1.3.36	R04150	2.51962
2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate	|2-Oxoglutarate| + |H+| + |Isochorismate| => |CO2| + |2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohex	2.2.1.9	R08165	2.51962
rxn04673	|1-4-Dihydroxy-2-naphthoate| + |Phytyl diphosphate| => |CO2| + |PPi| + |H+| + |Demethylphylloquin	2.5.1.-	R06858	0
1,4-dihydroxy-2-naphthoate octaprenyltransferase	|H+| + |1-4-Dihydroxy-2-naphthoate| + |FarnesylfarnesylgeraniolI| => |CO2| + |PPi| + |2-Demethylmer	Undetermined	None	2.51962
cpd00155 phosphorylase	|Phosphate| + |H+| + |Glycogen| <=> |Glucose-1-phosphate| + |glycogen(n-1)|	2.4.1.1	None	739.277
glycogen synthase (ADPGlc)	|ADPglucose| + |glycogen(n-1)| <=> |ADP| + |Glycogen|	2.4.1.21	None	739.277
ATP:alpha-D-glucose-1-phosphate adenyltransferase	|ATP| + |Glucose-1-phosphate| <=> |PPi| + |ADPglucose|	2.7.7.27	R00948	739.277
D-sorbitol transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |Sorbitol[e]| <=> |Pyruvate| + |Sorbitol 6-phosphate|	Undetermined	None	0
L-threonine ammonia-lyase	|L-Threonine| => |NH3| + |2-Oxobutyrate|	4.3.1.19	R00996	0
L-Alanine:NAD+ oxidoreductase (deaminating)	|H2O| + |NAD| + |L-Alanine| <=> |NADH| + |NH3| + |Pyruvate| + |H+|	1.4.1.1	R00396	999.908
L-serine ammonia-lyase	|L-Serine| => |NH3| + |Pyruvate|	4.2.1.13,4.3.1.15,4.3.1.17	R00220,R00223	0
3-Phosphoserine:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |phosphoserine| <=> |L-Glutamate| + |3-Phosphonooxypyruvate|	2.6.1.52	R04173	0
O-Phospho-4-hydroxy-L-threonine:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |4-(Phosphonooxy)-threonine| <=> |L-Glutamate| + |2-Oxo-3-hydroxy-4-phosphobu	2.6.1.52	R05085	0
alpha-D-Glucose 6-phosphate ketol-isomerase	|D-glucose-6-phosphate| <=> |beta-D-Glucose 6-phosphate|	5.1.3.15,5.3.1.9	R02739	-348.634
beta-D-Glucose 6-phosphate ketol-isomerase	|beta-D-Glucose 6-phosphate| <=> |D-fructose-6-phosphate|	5.3.1.9	R03321	-348.634
D-Glucose-6-phosphate ketol-isomerase	|D-glucose-6-phosphate| <=> |D-fructose-6-phosphate|	5.3.1.9	R02740,R00771	1000
rxn02527	|NADH| + |H+| + |Butanal| <=> |NAD| + |n-Butanol|	1.1.1.-	R03544	0
Biotin ABC transporter	|H2O| + |ATP| + |BIOT[e]| => |ADP| + |Phosphate| + |H+| + |BIOT|	Undetermined	None	0
PIt6	|Phosphate[e]| + |H+[e]| <= |Phosphate| + |H+|	TC-2.A.20,2.A.20	None	0
phosphate ABC transporter permease protein	|PPi[e]| + |H+[e]| <= |PPi| + |H+|	Undetermined	None	0
membrane alanyl aminopeptidase	|H2O| + |Cys-Gly| <=> |Glycine| + |L-Cysteine|	3.4.11.2,3.4.11.1,3.4.11.2	R00899	-1.25981
aminopeptidase	|H2O| + |ala-L-asp-L| <=> |L-Alanine| + |L-Aspartate|	Undetermined	None	0
aminopeptidase	|H2O| + |gly-glu-L| <=> |L-Glutamate| + |Glycine|	Undetermined	None	0
aminopeptidase	|H2O| + |H+| + |met-L-ala-L| <=> |L-Alanine| + |L-Methionine|	Undetermined	None	0
aminopeptidase	|H2O| + |gly-asp-L| <=> |Glycine| + |L-Aspartate|	Undetermined	None	0
aminopeptidase	|H2O| + |gly-pro-L| => |Glycine| + |H+| + |L-Proline|	Undetermined	None	0
aminopeptidase	|H2O| + |Ala-Gln| <=> |L-Alanine| + |L-Glutamine|	Undetermined	None	0
aminopeptidase	|H2O| + |ala-L-glu-L| <=> |L-Glutamate| + |L-Alanine|	Undetermined	None	0
aminopeptidase	|H2O| + |L-alanylglycine| <=> |Glycine| + |L-Alanine|	Undetermined	None	0
aminopeptidase	|H2O| + |Ala-Leu| => |L-Alanine| + |L-Leucine|	Undetermined	None	0
aminopeptidase	|H2O| + |Gly-Gln| <=> |Glycine| + |L-Glutamine|	Undetermined	None	0
Gly-Phe aminopeptidase	|H2O| + |Gly-Phe| <=> |Glycine| + |L-Phenylalanine|	3.4.11.2	None	0
Gly-Try aminopeptidase	|H2O| + |Gly-Tyr| <=> |Glycine| + |L-Tyrosine|	3.4.11.2	None	0
Gly-Cys aminopeptidase	|H2O| + |Gly-Cys| <=> |Glycine| + |L-Cysteine|	3.4.11.2	None	12.8165
aminopeptidase	|H2O| + |Ala-His| <=> |L-Alanine| + |L-Histidine|	Undetermined	None	13.003
aminopeptidase	|H2O| + |Gly-Met| <=> |Glycine| + |L-Methionine|	Undetermined	None	19.2689
aminopeptidase	|H2O| + |gly-asn-L| <=> |Glycine| + |L-Asparagine|	Undetermined	None	23.5354
aminopeptidase	|H2O| + |ala-L-Thr-L| <=> |L-Alanine| + |L-Threonine|	Undetermined	None	29.6468
Gly-Leu aminopeptidase	|H2O| + |Gly-Leu| <=> |Glycine| + |L-Leucine|	3.4.11.2	None	47.8096
NADH dehydrogenase (menaquinone-8 & 0 protons)	|NADH| + |H+| + |Menaquinone 8| => |NAD| + |Menaquinol 8|	1.6.5.3	None	0
NADH dehydrogenase (ubiquinone-8)	|NADH| + |H+| + |Ubiquinone-8| => |NAD| + |Ubiquinol-8|	1.6.5.3	None	0
NADH dehydrogenase (demethylmenaquinone-8 & 0 protons)	|NADH| + |H+| + |2-Demethylmenaquinone 8| => |NAD| + |2-Demethylmenaquinol 8|	1.6.5.3	None	0
LL-2,6-Diaminoheptanedioate 2-epimerase	|LL-2,6-Diaminopimelate| <=> |meso-2,6-Diaminopimelate|	5.1.1.7	R02735	0.286573
ATP:L-homoserine O-phosphotransferase	|ATP| + |L-Homoserine| <=> |ADP| + |O-Phospho-L-homoserine|	2.7.1.39	R01771	0
O-Phospho-L-homoserine phospho-lyase (adding water)	|H2O| + |O-Phospho-L-homoserine| => |Phosphate| + |H+| + |L-Threonine|	4.2.3.1,4.2.99.2	R01466	0
O-Phospho-4-hydroxy-L-threonine phospho-lyase (adding water)	|H2O| + |4-(Phosphonooxy)-threonine| => |Phosphate| + |H+| + |4-Hydroxy-L-threonine|	4.2.3.1	R05086	0
L-Homoserine:NADP+ oxidoreductase	|NADP| + |L-Homoserine| <=> |NADPH| + |H+| + |L-Aspartate4-semialdehyde|	1.1.1.3	R01775	-1000
L-Homoserine:NAD+ oxidoreductase	|NAD| + |L-Homoserine| <=> |NADH| + |H+| + |L-Aspartate4-semialdehyde|	1.1.1.3	R01773	1000
D-Glycerate:NADP+ 2-oxidoreductase	|NAD| + |Glycerate| <=> |NADH| + |H+| + |Hydroxypyruvate|	1.1.1.26,1.1.1.29,1.1.1.8	R01388	-1000
Glycolate:NADP+ oxidoreductase	|NADP| + |Glycolate| <=> |NADPH| + |Glyoxalate| + |H+|	1.1.1.79,1.1.1.26	R00465	-392.144
D-Glycerate:NADP+ 2-oxidoreductase	|NADP| + |Glycerate| <=> |NADPH| + |H+| + |Hydroxypyruvate|	1.1.1.79,1.1.1.81	R01392	0
Glycolate:NAD+ oxidoreductase	|NAD| + |Glycolate| <=> |NADH| + |Glyoxalate| + |H+|	1.1.1.26,1.1.1.29	R00717	1000
L-Leucine-ABC transport	|H2O| + |ATP| + |L-Leucine[e]| => |ADP| + |Phosphate| + |H+| + |L-Leucine|	Undetermined	None	0
L-Isoleucine-ABC transport	|H2O| + |ATP| + |L-Isoleucine[e]| => |ADP| + |Phosphate| + |H+| + |L-Isoleucine|	Undetermined	None	0
dihydrolipoylprotein:NAD+ oxidoreductase	|NAD| + |Dihydrolipolprotein| <=> |NADH| + |H+| + |Lipoylprotein|	1.8.1.4	R03815	0
S-aminomethyldihydrolipoylprotein:(6S)-tetrahydrofolate	|Tetrahydrofolate| + |S-Aminomethyldihydrolipoylprotein| => |NH3| + |5-10-Methylenetetrahydrofola	2.1.2.10	R04125	0
3-Hydroxy-2-methylpropanoate:NAD+ oxidoreductase	|NAD| + |3-Hydroxyisobutyrate| <=> |NADH| + |H+| + |3-Oxo-2-methylpropanoate|	1.1.1.31,1.1.1.35	R02047,R05066	0
(2S,3S)-3-Hydroxy-2-methylbutanoyl-CoA hydro-liase	|2-methyl-3-hydroxy-butyryl-CoA| <=> |H2O| + |Tiglyl-CoA|	4.2.1.17	R04204	0
trans-4-Hydroxy-L-proline:NAD+ oxidoreductase	|FAD| + |trans-4-Hydroxy-L-proline| <=> |FADH2| + |3-Hydroxy-L-1-pyrroline-5-carboxylate|	1.5.1.12,1.5.99.8,1.5.-.-	R03295	0
Proline dehydrogenase	|FAD| + |L-Proline| <=> |FADH2| + |1-Pyrroline-5-carboxylate|	1.5.99.8	None	1000
7,8-Diaminononanoate:carbon-dioxide cyclo-ligase	|ATP| + |CO2| + |7-8-Diaminononanoate| => |ADP| + |Phosphate| + (3) |H+| + |Dethiobiotin|	6.3.3.3	R03182	0
L-Valine-ABC transport	|H2O| + |ATP| + |L-Valine[e]| => |ADP| + |Phosphate| + |H+| + |L-Valine|	Undetermined	None	0
5-carboxymethyl-2-hydroxymuconic-semialdehyde:NAD+ oxidoreductase	|H2O| + |NAD| + |5-Carboxymethyl-2-hydroxymuconic semialdehyde| => |NADH| + (2) |H+| + |5-Carb	1.2.1.45,1.2.1.60	R04418	0
5-Carboxymethyl-2-hydroxymuconate	|5-Carboxymethyl-2-hydroxymuconate| <=> |5-Carboxy-2-oxohept-3-enedioate|	5.3.3.10	R04379	0
rxn02885	|2-Oxohept-3-enedioate| <= |2-Hydroxyhepta-2,4-dienedioate|	5.3.3.-	R04134	0
5-Oxopent-3-ene-1,2,5-tricarboxylate carboxy-lyase	|H+| + |5-Carboxy-2-oxohept-3-enedioate| => |CO2| + |2-Hydroxyhepta-2,4-dienedioate|	4.1.1.68	R04380	0
Uroporphyrinogen I carboxy-lyase	(4) |H+| + |UroporphyrinogenI| <=> (4) |CO2| + |CoproporphyrinogenI|	4.1.1.37	R04972	0
Uroporphyrinogen-III carboxy-lyase	(4) |H+| + |UroporphyrinogenIII| <=> (4) |CO2| + |CoproporphyrinogenIII|	4.1.1.37	R03197	1.25981
3,4-Dihydroxyphenylacetate:oxygen 2,3-oxidoreductase (decyclizing)	|O2| + |Homoprotocatechuate| => |H+| + |5-Carboxymethyl-2-hydroxymuconic semialdehyde|	1.13.11.15	R03303	0
4-hydroxyphenylacetate,NADH:oxygen oxidoreductase (3-hydroxylating)	|NADH| + |O2| + |H+| + |4-Hydroxyphenylacetate| => |H2O| + |NAD| + |Homoprotocatechuate|	1.14.13.3	R02698	0
3-Hydroxyphenylacetate,NADH:oxygen oxidoreductase (3-hydroxylating)	|NADH| + |O2| + |H+| + |3-Hydroxyphenylacetate| => |H2O| + |NAD| + |Homoprotocatechuate|	1.14.13.3	R03299	0
hydrogen-peroxide:hydrogen-peroxide oxidoreductase	(2) |H2O2| => (2) |H2O| + |O2|	1.11.1.6	R00009	0
Protoporphyrin ferro-lyase	|Protoporphyrin| + |Fe2+| <=> |Heme| + (2) |H+|	4.99.1.1	R00310	1.25981
Protoporphyrinogen-IX:oxygen oxidoreductase	(3) |O2| + (2) |ProtoporphyrinogenIX| <=> (6) |H2O| + (2) |Protoporphyrin|	1.3.3.4	R03222	0.629905
2-Phospho-D-glycerate hydro-lyase	|2-Phospho-D-glycerate| <=> |H2O| + |Phosphoenolpyruvate|	4.2.1.11	R00658	1000
D-Glyceraldehyde-3-phosphate ketol-isomerase	|Glyceraldehyde3-phosphate| <=> |Glycerone-phosphate|	5.3.1.1	R01015	-1000
ATP:3-phospho-D-glycerate 1-phosphotransferase	|ATP| + |H+| + |3-Phosphoglycerate| <=> |ADP| + |1,3-Bisphospho-D-glycerate|	2.7.2.3	R01512	-1000
D-erythrose 4-phosphate:NAD+ oxidoreductase	|H2O| + |NAD| + |D-Erythrose4-phosphate| => |NADH| + (2) |H+| + |4-Phosphoerythronate|	1.2.1.72	R01825	0
D-Glyceraldehyde-3-phosphate:NAD+ oxidoreductase(phosphorylating)	|NAD| + |Phosphate| + |Glyceraldehyde3-phosphate| <=> |NADH| + |1,3-Bisphospho-D-glycerate|	1.2.1.12,1.2.1.13,1.2.1.59	R01061	1000
Phosphoribosyl-ATP pyrophosphohydrolase	|H2O| + |Phosphoribosyl-ATP| => |PPi| + |Phosphoribosyl-AMP|	3.6.1.31	R04035	0
1-(5-phospho-D-ribosyl)-AMP 1,6-hydrolase	|H2O| + |Phosphoribosyl-AMP| <=> |phosphoribosylformiminoaicar-phosphate|	3.5.4.19	R04037	0
Imidazole-glycerol-3-phosphate synthase	|L-Glutamate| + |H+| + |D-erythro-imidazol-glycerol-phosphate| + |AICAR| <= |L-Glutamine| + |phosp	Undetermined	R04558	0
N-(5'-Phospho-D-ribosylformimino)-5-amino-1-	|phosphoribosylformiminoaicar-phosphate| <=> |phosphoribulosylformimino-AICAR-phosphate|	5.3.1.16	R04640	0
D-erythro-1-(Imidazol-4-yl)glycerol 3-phosphate hydro-lyase	|D-erythro-imidazol-glycerol-phosphate| => |H2O| + |imidazole acetol-phosphate|	4.2.1.19	R03457	0
L-Histidinal:NAD+ oxidoreductase	|H2O| + |NAD| + |L-Histidinal| => |NADH| + (2) |H+| + |L-Histidine|	1.1.1.23	R01163	0
L-Histidinol:NAD+ oxidoreductase	|NAD| + |L-Histidinol| <=> |NADH| + |H+| + |L-Histidinal|	1.1.1.23	R03012	0
1-(5-Phospho-D-ribosyl)-ATP:pyrophosphate phosphoribosyl-transferase	|PPi| + |Phosphoribosyl-ATP| <= |ATP| + |PRPP|	2.4.2.17	R01071	0
UDP-D-galactopyranose furanomutase	|UDP-galactose| <=> |UDP-D-galacto-1,4-furanose|	5.4.99.9	R00505	0
D-Alanine:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |D-Alanine| <=> |Pyruvate| + |D-Glutamate|	2.6.1.21	R01148	-999.713
UDP-N-acetylmuramate:NADP+ oxidoreductase	|NADP| + |UDP-MurNAc| <=> |NADPH| + |H+| + |UDP-N-acetylglucosamine enolpyruvate|	1.1.1.158	R03192	-0.286573
palmitoyl-lipoteichoic acid synthesis (n=24), linked, N-acetylglucosamine sub	(24) |UDP-N-acetylglucosamine| + |Palmitoyllipoteichoic acid (n=24), linked, unsubstituted| <=> (24) |U	2.7.8.-	None	0
myristoyl-lipoteichoic acid synthesis (n=24), linked, N-acetylglucosamine sub	(24) |UDP-N-acetylglucosamine| + |Myristoyllipoteichoic acid (n=24), linked, unsubstituted| <=> (24) |U	2.7.8.-	None	0
isotetradecanoyl-lipoteichoic acid synthesis (n=24), linked, N-acetylglucosam	(24) |UDP-N-acetylglucosamine| + |Isotetradecanoyllipoteichoic acid (n=24), linked, unsubstituted| <=>	2.7.8.-	None	0
isopentadecanoyl-lipoteichoic acid synthesis (n=24), linked, N-acetylglucosar	(24) |UDP-N-acetylglucosamine| + |Isopentadecanoyllipoteichoic acid (n=24), linked, unsubstituted| <=	2.7.8.-	None	0
anteisopentadecanoyl-lipoteichoic acid synthesis (n=24), linked, N-acetylgluc	(24) |UDP-N-acetylglucosamine| + |Anteisopentadecanoyllipoteichoic acid (n=24), linked, unsubstituted	2.7.8.-	None	0
isohexadecanoyl-lipoteichoic acid synthesis (n=24), linked, N-acetylglucosam	(24) |UDP-N-acetylglucosamine| + |Isohexadecanoyllipoteichoic acid (n=24), linked, unsubstituted| <=>	2.7.8.-	None	0

Reactions with corresponding equation, EC number and FBA flux value under aerobic conditions in rich media

stearoyl-lipoteichoic acid synthesis (n=24), linked, N-acetylglucosamine subs	(24) |UDP-N-acetylglucosamine| + |Stearoyllipoteichoic acid (n=24), linked, unsubstituted| <=> (24) |UI	2.7.8.-	None	0.286573
isoheptadecanoyl-lipoteichoic acid synthesis (n=24), linked, N-acetylglucosar	(24) |UDP-N-acetylglucosamine| + |Isoheptadecanoyllipoteichoic acid (n=24), linked, unsubstituted| <=	2.7.8.-	None	0.286573
anteisoheptadecanoyl-lipoteichoic acid synthesis (n=24), linked, N-acetylgluc	(24) |UDP-N-acetylglucosamine| + |Anteisoheptadecanoyllipoteichoic acid (n=24), linked, unsubstituted|	2.7.8.-	None	0.286573
UDP-N-acetylglucosamine:undecaprenylphosphate N-acetylglucosamine -1-p	|UDP-N-acetylglucosamine| + |Undecaprenylphosphate| <=> |UMP| + |Undecaprenyl diphospho N-ace	Undetermined	R08856	0.859719
Isocitrate glyoxylate-lyase	|Isocitrate| <=> |Succinate| + |Glyoxalate|	4.1.3.1	R00479	0
UDP-N-acetyl-D-mannosamine:N-acetyl-beta-D-	|UDP-N-acetyl-D-mannosamine| + |Undecaprenyl diphospho N-acetyl-glucosamine| <=> |UDP| + |N-Ac	2.4.1.187	R05566	0.859719
UDPglucose:NAD+ oxidoreductase	|H2O| + (2) |NAD| + |UDP-glucose| <=> (2) |NADH| + (3) |H+| + |UDPglucuronate|	1.1.1.22	R00286	0
UDP-N-acetyl-D-glucosamine 2-epimerase	|H2O| + |UDP-N-acetylglucosamine| => |UDP| + |N-Acetyl-D-mannosamine|	5.1.3.14	R00414	0
UDP-N-acetyl-D-glucosamine 2-epimerase	|UDP-N-acetylglucosamine| <=> |UDP-N-acetyl-D-mannosamine|	5.1.3.14	R00420	0.859719
ATP:D-fructose 6-phosphotransferase	|ATP| + |D-Fructose| <=> |ADP| + |D-fructose-6-phosphate|	2.7.1.1,2.7.1.4	R00760,R00867,R0	-167.762
FMNH2:NAD+ oxidoreductase	|NAD| + |FMNH2| <= |NADH| + |FMN| + |H+|	1.5.1.29	R05705	0
FMNH2-dependent monooxygenase	|O2| + |FMNH2| + |Isethionate| => |H2O| + |FMN| + |Sulfite| + |Glycolaldehyde|	Undetermined	None	0
FMNH2-dependent monooxygenase (methanesulfonate)	|O2| + |FMNH2| + |methanesulfonate| => |H2O| + |FMN| + |Formaldehyde| + |Sulfite|	Undetermined	None	0
FMNH2-dependent monooxygenase (ethanesulfonate)	|O2| + |FMNH2| + |ethanesulfonate| => |H2O| + |FMN| + |Acetaldehyde| + |Sulfite|	Undetermined	None	0
FMNH2-dependent monooxygenase (butanesulfonate)	|O2| + |FMNH2| + |butanesulfonate| => |H2O| + |FMN| + |Sulfite| + |Butanal|	Undetermined	None	0
FMNH2-dependent monooxygenase (sulfoacetate)	|O2| + |FMNH2| + |Sulfoacetate| => |H2O| + |Glyoxalate| + |FMN| + |Sulfite|	Undetermined	None	0
FMN-dependent monooxygenase (hexanesulfonate)	|O2| + |FMNH2| + |hexanesulfonate| => |H2O| + |FMN| + |Sulfite| + |hexanal|	Undetermined	None	0
FMN-dependent monooxygenase (taurine)	|O2| + |Taurine| + |FMNH2| => |H2O| + |FMN| + |Sulfite| + |Aminoacetaldehyde|	Undetermined	None	1.25981
Arbutin 6-phosphate glucohydrolase	|H2O| + |Arbutin-6P| => |Quinol| + |beta-D-Glucose 6-phosphate|	3.2.1.86	R05133	0
Salicin 6-phosphate glucohydrolase	|H2O| + |Salicin-6P| => |beta-D-Glucose 6-phosphate| + |Saligenin|	3.2.1.86	R05134	0
Arbutin 6-phosphate glucohydrolase	|H2O| + |Arbutin-6P| <=> |D-glucose-6-phosphate| + |Quinol|	3.2.1.86	None	0
Salicin 6-phosphate glucohydrolase	|H2O| + |Salicin-6P| => |D-glucose-6-phosphate| + |Saligenin|	3.2.1.86	None	0
6-Phospho-beta-D-glucosyl-(1,4)-D-gluxose glucohydrolase	|H2O| + |cellobiose 6-phoshate| => |D-Glucose| + |D-glucose-6-phosphate|	3.2.1.86	R00839	832.238
ATP:D-ribose 5-phosphotransferase	|ATP| + |D-Ribose| <=> |ADP| + |ribose-5-phosphate|	2.7.1.15	R01051	0
Deoxyribokinase	|ADP| + |deoxyribose-5-phosphate| <=> |ATP| + |Thyminose|	2.7.1.15	R02750	0
O-Acetyl-L-homoserine acetate-lyase (adding methanethiol)	|H2S| + |O-Acetyl-L-homoserine| => |Acetate| + |H+| + |Homocysteine|	2.5.1.49,4.2.99.8,4.2.99.2	R01287	0
O-Succinyl-L-homoserine succinate-lyase (adding cysteine)	|H2S| + |O-Succinyl-L-homoserine| => |Succinate| + |H+| + |Homocysteine|	2.5.1.48,4.2.99.9,2.5.1.-	R01288	0
O-Acetyl-L-homoserine acetate-lyase (adding methanethiol)	|H2S2O3| + |O-Acetyl-L-homoserine| + |trdrd| => |Acetate| + |Sulfite| + |Homocysteine| + |trdox|	2.5.1.49	R02026	0
Chorismate hydroxymutase	|Chorismate| <=> |Isochorismate|	5.4.4.2,5.4.99.6	R01717	2.51962
L-Glutamine:D-fructose-6-phosphate aminotransferase (hexose	|L-Glutamine| + |D-fructose-6-phosphate| <=> |L-Glutamate| + |D-Glucosamine phosphate|	2.6.1.16	R00768	665.844
L-Arginine amidinohydrolase	|H2O| + |L-Arginine| => |Ornithine| + |Urea|	3.5.3.1	R00551	969.313
adenylate kinase (Inorganic triphosphate)	|AMP| + |H+| + |Triphosphate| <=> |ADP| + |PPi|	2.7.4.3	None	-998.74
ATP:AMP phosphotransferase	|ATP| + |dAMP| <=> |ADP| + |dADP|	2.7.4.11,2.7.4.3	R01547	2.13204
ATP:AMP phosphotransferase	|ATP| + |AMP| <=> (2) |ADP|	2.7.4.3	R00127	1000
L-Ornithine:2-oxo-acid aminotransferase	|2-Oxoglutarate| + |Ornithine| <=> |L-Glutamate| + |L-Glutamate5-semialdehyde|	2.6.1.13	R00667	0
Protein biosynthesis	=> |Protein biosynthesis|	Undetermined	None	158.673
serine O-acetyltransferase	|Acetyl-CoA| + |L-Serine| => |CoA| + |O-Acetyl-L-serine|	2.3.1.30	R00586	0
L-Glutamate:tRNA(Glu) ligase (AMP-forming)	|ATP| + |L-Glutamate| + |H+| + |tRNA-Glu| => |PPi| + |AMP| + |L-Glutamyl-tRNA-Glu|	6.1.1.17,6.1.1.24	R05578	20.157
2-Phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol	|2-phospho-4--cytidine5-diphospho-2-C-methyl-D-erythritol| <=> |CMP| + |2-C-methyl-D-erythritol2-4-	4.6.1.12	R05637	33.3877
CTP: 2-C-Methyl-D-erythritol 4-phosphate cytidylyltransferase	|CTP| + |2-C-methyl-D-erythritol4-phosphate| <=> |PPi| + |4--cytidine5-diphospho-2-C-methyl-D-eryth	2.7.7.60	R05633	33.3877
ATP:2-amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine	|ATP| + |6-hydroxymethyl dihydropterin| <=> |AMP| + |7,8-Dihydropterin pyrophosphate|	2.7.6.3	R03503	0
2-Amino-4-hydroxy-6-(D-erythro-1,2,3-trihydroxypropyl)-7,8-	|Dihydroneopterin| <=> |Glycolaldehyde| + |6-hydroxymethyl dihydropterin|	4.1.2.25	R03504	3.77943
L-aspartate:L-glutamine amido-ligase (AMP-forming)	|H2O| + |ATP| + |L-Aspartate| + |L-Glutamine| => |PPi| + |AMP| + |L-Glutamate| + |L-Asparagine|	6.3.5.4	R00578	0
2-Amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine-	|ABEE| + |7,8-Dihydropterin pyrophosphate| => |PPi| + |Dihydropteroate|	2.5.1.15	R03067	0
2-Amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine:4-	|ABEE| + |6-hydroxymethyl dihydropterin| <=> |H2O| + |Dihydropteroate|	2.5.1.15	R03066	3.77943
4-amino-4-deoxychorismate pyruvate-lyase	|ADC| => |Pyruvate| + |H+| + |ABEE|	4.1.3.38	R05553	3.77943
Chorismate pyruvate-lyase (amino-accepting)	|NH3| + |Chorismate| => |H2O| + |Pyruvate| + |H+| + |Anthranilate|	4.1.3.27	R00985	0
Chorismate pyruvate-lyase (amino-accepting)	|L-Glutamine| + |Chorismate| => |Pyruvate| + |L-Glutamate| + |H+| + |Anthranilate|	4.1.3.27	R00986	0
chorismate:L-glutamine aminotransferase	|L-Glutamine| + |Chorismate| <=> |L-Glutamate| + |ADC|	2.6.1.85	R01716	3.77943
ATP:pantothenate 4'-phosphotransferase	|ATP| + |PAN| <=> |ADP| + |4-phosphopantothenate|	2.7.1.33	R03018	2.51962
IMP:pyrophosphate phosphoribosyltransferase	|PPi| + |IMP| <= |PRPP| + |HYXN|	2.4.2.8	R01132	-27.788
GMP:pyrophosphate phosphoribosyltransferase	|PPi| + |GMP| <= |PRPP| + |Guanine|	2.4.2.7,2.4.2.8,2.4.2.22	R01229	0
ATP:D-ribose-5-phosphate pyrophosphotransferase	|ATP| + |ribose-5-phosphate| <=> |AMP| + |PRPP|	2.7.6.1	R01049	36.9943
UTP:N-acetyl-alpha-D-glucosamine-1-phosphate uridylyltransferase	|UTP| + |N-Acetyl-D-glucosamine1-phosphate| <=> |PPi| + |UDP-N-acetylglucosamine|	2.7.7.23	R00416	31.523
Acetyl-CoA:D-glucosamine-1-phosphate N-acetyltransferase	|Acetyl-CoA| + |D-Glucosamine1-phosphate| => |CoA| + |H+| + |N-Acetyl-D-glucosamine1-phosphate|	2.3.1.157,2.3.1.4	R05332	31.523
ATP: 4-(Cytidine 5'-diphospho)-2-C-methyl-D-erythritol	|ATP| + |4--cytidine5-diphospho-2-C-methyl-D-erythritol| <=> |ADP| + |2-phospho-4--cytidine5-diphos	2.7.1.148	R05634	33.3877
DNA replication	=> |DNA replication|	Undetermined	None	158.673
ATP:dTMP phosphotransferase	|ATP| + |dTMP| <=> |ADP| + |dTDP|	2.7.4.12,2.7.4.9	R02094	2.13204
ATP:deoxyguanosine 5'-phosphotransferase	|ATP| + |Deoxyguanosine| <=> |ADP| + |dGMP|	2.7.1.113	R01967	-1000
ATP:deoxyadenosine 5'-phosphotransferase	|ATP| + |Deoxyadenosine| <=> |ADP| + |dAMP|	2.7.1.76	R02089	2.13204
rxn05144	|L-Glutamine| + |Glyceraldehyde3-phosphate| + |D-Ribulose5-phosphate| => (3) |H2O| + |Phosphate|	Undetermined	None	1.25981
Potassium uptake	|K+[e]| <=> |K+|	1.A.1	None	-997.48
Formate:NAD+ oxidoreductase	|NAD| + |Formate| => |NADH| + |CO2|	1.2.1.2	R00519	0
IMP:L-aspartate ligase (GDP-forming)	|GTP| + |L-Aspartate| + |IMP| => |Phosphate| + |GDP| + (2) |H+| + |Adenylosuccinate|	6.3.4.4	R01135	15.1484
6-Phospho-D-gluconate:NADP+ 2-oxidoreductase (decarboxylating)	|NADP| + |6-Phospho-D-gluconate| => |NADPH| + |CO2| + |D-Ribulose5-phosphate|	1.1.1.44	R01528	0
ATP:(R)-glycerate 3-phosphotransferase	|ATP| + |Glycerate| <=> |ADP| + |3-Phosphoglycerate|	2.7.1.31	R01514	0
D-Mannose-6-phosphate ketol-isomerase	|D-mannose-6-phosphate| <=> |D-fructose-6-phosphate|	5.3.1.8	R01819,R00772	-1000
Acetyl-CoA:formate C-acetyltransferase	|Acetyl-CoA| + |Formate| <=> |CoA| + |Pyruvate|	2.3.1.54	R00212	-1000
Acetyl-CoA:formate C-acetyltransferase	|CoA| + |2-Oxobutyrate| <=> |Formate| + |Propionyl-CoA|	2.3.1.54	R06987	0
menauinol oxidase (7:2 protons)	(0.5) |O2| + (4) |H+| + |mql7| => |H2O| + (4) |H+[e]| + |Menaquinone 7|	Undetermined,,Undeterr	None	0
Na+:proline symport	|L-Proline[e]| + |Na+[e]| <=> |L-Proline| + |Na+|	Undetermined	None	1000
(R)-Malate:NAD+ oxidoreductase (decarboxylating)	|NAD| + |D-Malate| => |NADH| + |CO2| + |Pyruvate|	1.1.1.83	R00215	0
(R,R)-Tartrate carboxy-lyase	|H+| + |Tartrate| => |CO2| + |Glycerate|	4.1.1.73	R01751	0
meso-Tartaric acid:NAD+ oxidoreductase	|NAD| + |meso-Tartrate| <=> |NADH| + |H+| + |Oxaloglycolate|	1.1.1.93	R02545	0
tartrate dehydrogenase	|NAD| + |Tartrate| <=> |NADH| + |H+| + |Oxaloglycolate|	1.1.1.93	R06180	0
Acetyl-CoA:orthophosphate acetyltransferase	|Phosphate| + |Acetyl-CoA| + |H+| <=> |CoA| + |Acetylphosphate|	2.3.1.8	R00230	865.74
acetyl-CoA:phosphate acetyltransferase	|Phosphate| + |H+| + |Propionyl-CoA| <=> |CoA| + |Propionyl phosphate|	2.3.1.8	R00921	1000
ATP:lipoate adenylyltransferase	|ATP| + |H+| + |Lipoate| <=> |PPi| + |Lipoyl-AMP|	2.7.7.63	R07770	0
dGTP triphosphohydrolase	|H2O| + |dGTP| <=> |H+| + |Deoxyguanosine| + |Triphosphate|	3.1.5.1	R01856	-1000
Hypoxanthine ion-coupled transport	|H+[e]| + |HYXN[e]| <=> |H+| + |HYXN|	Undetermined	None	-972.212
4-oxalocrotonate tautomerase	|2-Hydroxymuconate| => |4-Oxalocrotonate|	5.3.2.-	R03966	0
rxn03644	|D-arabino-6-Phospho-hex-3-ulose| <=> |D-fructose-6-phosphate|	5.-.-.-	R05339	0
S-Adenosylmethioninamine:putrescine 3-aminopropyltransferase	|Putrescine| + |S-Adenosylmethioninamine| => |H+| + |5-Methylthioadenosine| + |Spermidine|	2.5.1.16	R01920	0
Agmatine amidinohydrolase	|H2O| + |Agmatine| => |Urea| + |Putrescine|	3.5.3.11	R01157	0
D-arabino-3-Hexulose 6-phosphate formaldehyde-lyase	|Formaldehyde| + |D-Ribulose5-phosphate| <=> |D-arabino-6-Phospho-hex-3-ulose|	4.1.2.-	R05338	0
(R)-2-Methyl-3-oxopropanoyl-CoA CoA-carbonylmutase	|L-methylmalonyl-CoA| <=> |Succinyl-CoA|	5.4.99.2	R00833	0
UTP:ammonia ligase(ADP-forming)	|ATP| + |NH3| + |UTP| <=> |ADP| + |Phosphate| + |CTP| + (2) |H+|	6.3.4.2	R00571	-822.764
UTP:ammonia ligase(ADP-forming)	|H2O| + |ATP| + |L-Glutamine| + |UTP| => |ADP| + |Phosphate| + |L-Glutamate| + |CTP| + (2) |H+|	6.3.4.2	R00573	0
Sedoheptulose 1,7-bisphosphate D-glyceraldehyde-3-phosphate-lyase	|Sedoheptulose 1,7-bisphosphate| <=> |Glycerone-phosphate| + |D-Erythrose4-phosphate|	4.1.2.13	R01829	-1000
D-Fructose 1-phosphate D-glyceraldehyde-3-phosphate-lyase	|D-fructose-1-phosphate| <=> |Glycerone-phosphate| + |D-Glyceraldehyde|	4.1.2.13	R02568	729.838
D-Fructose-1,6-bisphosphate D-glyceraldehyde-3-phosphate-lyase	|D-fructose-1,6-bisphosphate| <=> |Glycerone-phosphate| + |Glyceraldehyde3-phosphate|	4.1.2.13	R01068,R01070	828.266
Sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate	|Glyceraldehyde3-phosphate| + |Sedoheptulose7-phosphate| <=> |D-fructose-6-phosphate| + |D-Erytl	2.2.1.2	R01827,R08575	873.952
Phosphoenolpyruvate:UDP-N-acetyl-D-glucosamine	|UDP-N-acetylglucosamine| + |Phosphoenolpyruvate| <=> |Phosphate| + |H+| + |UDP-N-acetylglucosa	2.5.1.7	R00660	0.286573
D-Fructose-1,6-bisphosphate 1-phosphohydrolase	|H2O| + |D-fructose-1,6-bisphosphate| => |Phosphate| + |H+| + |D-fructose-6-phosphate|	3.1.3.11	R00762,R04780	0
ATP:deoxyuridine 5'-phosphotransferase	|ATP| + |Deoxyuridine| <=> |ADP| + |dUMP|	2.7.1.21	R02099	-342.12
ATP:thymidine 5'-phosphotransferase	|ATP| + |Thymidine| <=> |ADP| + |dTMP|	2.7.1.21	R01567	0
(R,R)-Butane-2,3-diol:NAD+ oxidoreductase	|NAD| + |BDOH| <=> |NADH| + |H+| + |ACTN|	1.1.1.4	R02946	0
(S,S)-Butane-2,3-diol:NAD+ oxidoreductase	|NAD| + |(S,S)-2,3-Butanediol| <=> |NADH| + |H+| + |(S)-Acetoin|	1.1.1.76	R03707	0
D-Ribose-5-phosphate ketol-isomerase	|ribose-5-phosphate| <=> |D-Ribulose5-phosphate|	5.3.1.6	R01056	-258.395
5,10-Methylenetetrahydrofolate:glycine hydroxymethyltransferase	|H2O| + |Glycine| + |5-10-Methylenetetrahydrofolate| <=> |L-Serine| + |Tetrahydrofolate|	2.1.2.1	R00945	38.4805
UMP:pyrophosphate phosphoribosyltransferase	|PPi| + |UMP| <= |Uracil| + |PRPP|	2.4.2.9	R00966	0
ATP synthase (four protons for one ATP)	|ADP| + |Phosphate| + (4) |H+[e]| <=> |H2O| + |ATP| + (3) |H+|	3.6.3.14	None	1000
NADH dehydrogenase (ubiquinone-8 & 3.5 protons)	|NADH| + (4.5) |H+| + |Ubiquinone-8| <=> |NAD| + (3.5) |H+[e]| + |Ubiquinol-8|	1.6.5.3	None	1000
5-Methyltetrahydrofolate:L-homocysteine S-methyltransferase	|Homocysteine| + |5-Methyltetrahydrofolate| <=> |L-Methionine| + |Tetrahydrofolate|	2.1.1.13,2.1.1.14	R00946	7.55886
5-methyltetrahydrofolate:NADP+ oxidoreductase	|NADP| + |5-Methyltetrahydrofolate| <=> |NADPH| + |H+| + |5-10-Methylenetetrahydrofolate|	1.5.1.20,1.7.99.5	R01224	-138.789
5-methyltetrahydrofolate:NAD+ oxidoreductase	|NAD| + |5-Methyltetrahydrofolate| <=> |NADH| + |H+| + |5-10-Methylenetetrahydrofolate|	1.5.1.20	R07168	129.97
glutathione hydralase (periplasmic)	|H2O| + |GSH| <=> |L-Glutamate| + |Cys-Gly|	2.3.2.2,3.4.11.4	R00494	-1.25981
(5-Glutamyl)-peptide:amino-acid 5-glutamyltransferase	|L-Glutamate| + |L-3-Cyanoalanine| <=> |H2O| + |gamma-Glutamyl-beta-cyanoalanine|	2.3.2.2	R03970	0
(5-Glutamyl)-peptide:amino-acid 5-glutamyltransferase	|L-Glutamate| + |H+| + |L-3-Cyanoalanine| => |H2O| + |CO2| + |gamma-Glutamyl-3-aminopropiononi	2.3.2.2	None	0
4-methyl-3-hydroxy-pentanoyl-ACP hydro-lyase	|4-methyl-3-hydroxy-pentanoyl-ACP| <=> |H2O| + |4-methyl-trans-pent-2-enoyl-ACP|	4.2.1.0	None	0
6-methyl-3-hydroxy-heptanoyl-ACP hydro-lyase	|6-methyl-3-hydroxy-heptanoyl-ACP| <=> |H2O| + |6-methyl-trans-hept-2-enoyl-ACP|	4.2.1.0	None	0
8-methyl-3-hydroxy-nonanoyl-ACP hydro-lyase	|8-methyl-3-hydroxy-nonanoyl-ACP| <=> |H2O| + |8-methyl-trans-non-2-enoyl-ACP|	4.2.1.0	None	0
10-methyl-3-hydroxy-undecanoyl-ACP hydro-lyase	|10-methyl-3-hydroxy-undecanoyl-ACP| <=> |H2O| + |10-methyl-trans-undec-2-enoyl-ACP|	4.2.1.0	None	0
12-methyl-3-hydroxy-tridecanoyl-ACP hydro-lyase	|12-methyl-3-hydroxy-tridecanoyl-ACP| <=> |H2O| + |12-methyl-trans-tridec-2-enoyl-ACP|	4.2.1.0	None	0
14-methyl-3-hydroxy-pentadecanoyl-ACP hydro-lyase	|14-methyl-3-hydroxy-pentadecanoyl-ACP| <=> |H2O| + |14-methyl-trans-pentadec-2-enoyl-ACP|	4.2.1.0	None	0
(3R)-3-Hydroxybutanoyl-[acyl-carrier-protein] hydro-lyase	|(R)-3-Hydroxybutanoyl-[acyl-carrier protein]| <=> |H2O| + |But-2-enoyl-[acyl-carrier protein]|	4.2.1.58	R04428	13.193
(3R)-3-Hydroxybutanoyl-[acyl-carrier-protein] hydro-lyase	|D-3-Hydroxyhexanoyl-[acp]| <=> |H2O| + |(2E)-Hexenoyl-[acp]|	2.3.1.85,2.3.1.86,4.2.1.5	R04954	13.193
(3R)-3-Hydroxybutanoyl-[acyl-carrier-protein] hydro-lyase	|D-3-Hydroxydodecanoyl-[acp]| <=> |H2O| + |(2E)-Dodecenoyl-[acp]|	2.3.1.85,2.3.1.86,4.2.1.5	R04965	13.193

Reactions with corresponding equation, EC number and FBA flux value under aerobic conditions in rich media

(3R)-3-Hydroxypalmitoyl-[acyl-carrier-protein] hydro-lyase	|R-3-hydroxypalmitoyl-acyl-carrierprotein-| <=> |H2O| + |(2E)-Hexadecenoyl-[acp]|	2.3.1.85,2.3.1.86,4.2.1.6	R04544	13.193
(3R)-3-Hydroxybutanoyl-[acyl-carrier-protein] hydro-lyase	|(R)-3-Hydroxydecanoyl-[acyl-carrier protein]| <=> |H2O| + |(2E)-Decenoyl-[acp]|	2.3.1.85,2.3.1.86,4.2.1.5	R04535	13.193
(3R)-3-Hydroxybutanoyl-[acyl-carrier-protein] hydro-lyase	|(R)-3-Hydroxyoctanoyl-[acyl-carrier protein]| <=> |H2O| + |(2E)-Octenoyl-[acp]|	2.3.1.85,2.3.1.86,4.2.1.5	R04537	13.193
(3R)-3-Hydroxypalmitoyl-[acyl-carrier-protein] hydro-lyase	|HMA| <=> |H2O| + |(2E)-Tetradecenoyl-[acp]|	2.3.1.85,2.3.1.86,4.2.1.6	R04568	13.193
4-methyl-3-hydroxy-hexanoyl-ACP hydro-lyase	|4-methyl-3-hydroxy-hexanoyl-ACP| <=> |H2O| + |4-methyl-trans-hex-2-enoyl-ACP|	4.2.1.0	None	13.193
6-methyl-3-hydroxy-octanoyl-ACP hydro-lyase	|6-methyl-3-hydroxy-octanoyl-ACP| <=> |H2O| + |6-methyl-trans-oct-2-enoyl-ACP|	4.2.1.0	None	13.193
8-methyl-3-hydroxy-decanoyl-ACP hydro-lyase	|8-methyl-3-hydroxy-decanoyl-ACP| <=> |H2O| + |8-methyl-trans-dec-2-enoyl-ACP|	4.2.1.0	None	13.193
10-methyl-3-hydroxy-dodecanoyl-ACP hydro-lyase	|10-methyl-3-hydroxy-dodecanoyl-ACP| <=> |H2O| + |10-methyl-trans-dodec-2-enoyl-ACP|	4.2.1.0	None	13.193
12-methyl-3-hydroxy-tetra-decanoyl-ACP hydro-lyase	|12-methyl-3-hydroxy-tetra-decanoyl-ACP| <=> |H2O| + |12-methyl-trans-tetra-dec-2-enoyl-ACP|	4.2.1.0	None	13.193
14-methyl-3-hydroxy-hexa-decanoyl-ACP hydro-lyase	|14-methyl-3-hydroxy-hexa-decanoyl-ACP| <=> |H2O| + |14-methyl-trans-hexa-dec-2-enoyl-ACP|	4.2.1.0	None	13.193
5-methyl-3-hydroxy-hexanoyl-ACP hydro-lyase	|5-methyl-3-hydroxy-hexanoyl-ACP| <=> |H2O| + |5-methyl-trans-hex-2-enoyl-ACP|	4.2.1.0	None	13.193
7-methyl-3-hydroxy-octanoyl-ACP hydro-lyase	|7-methyl-3-hydroxy-octanoyl-ACP| <=> |H2O| + |7-methyl-trans-oct-2-enoyl-ACP|	4.2.1.0	None	13.193
9-methyl-3-hydroxy-decanoyl-ACP hydro-lyase	|9-methyl-3-hydroxy-decanoyl-ACP| <=> |H2O| + |9-methyl-trans-dec-2-enoyl-ACP|	4.2.1.0	None	13.193
11-methyl-3-hydroxy-dodecanoyl-ACP hydro-lyase	|11-methyl-3-hydroxy-dodecanoyl-ACP| <=> |H2O| + |11-methyl-trans-dodec-2-enoyl-ACP|	4.2.1.0	None	13.193
13-methyl-3-hydroxy-tetra-decanoyl-ACP hydro-lyase	|13-methyl-3-hydroxy-tetra-decanoyl-ACP| <=> |H2O| + |13-methyl-trans-tetra-dec-2-enoyl-ACP|	4.2.1.0	None	13.193
15-methyl-3-hydroxy-hexa-decanoyl-ACP hydro-lyase	|15-methyl-3-hydroxy-hexa-decanoyl-ACP| <=> |H2O| + |15-methyl-trans-hexa-dec-2-enoyl-ACP|	4.2.1.0	None	13.193
trans-Octodec-2-enoyl-ACP hydro-lyase	|3-Hydroxyoctodecanoyl-ACP| <=> |H2O| + |trans-Octodec-2-enoyl-ACP|	4.2.1.0	None	13.193
UTP:alpha-D-glucose-1-phosphate uridylyltransferase	|UTP| + |Glucose-1-phosphate| <=> |PPi| + |UDP-glucose|	2.7.7.9	R00289	0
D-Glucose-ABC transport	|H2O| + |ATP| + |D-Glucose[e]| => |ADP| + |Phosphate| + |D-Glucose| + |H+|	Undetermined	None	0
Branched chain amino acid:H+ symporter (Leucine)	|H+[e]| + |L-Leucine[e]| <=> |H+| + |L-Leucine|	Undetermined	None	20.516
L-valine transport in via proton symport	|H+[e]| + |L-Valine[e]| <=> |H+| + |L-Valine|	Undetermined	None	23.7121
Branched chain amino acid:H+ symporter (Isoleucine)	|H+[e]| + |L-Isoleucine[e]| <=> |H+| + |L-Isoleucine|	Undetermined	None	56.1028
Dimethylallyl-diphosphate:isopentenyl-diphosphate	|Isopentenyldiphosphate| + |DMAPP| => |PPi| + |Geranyldiphosphate|	2.5.1.1,2.5.1.10,2.5.1.29	R01658	4.066
Geranyl-diphosphate:isopentenyl-diphosphate geranyltrans-transferase	|Isopentenyldiphosphate| + |Geranyldiphosphate| => |PPi| + |Farnesyldiphosphate|	2.5.1.1,2.5.1.10,2.5.1.29	R02003	4.066
UDP-N-acetylglucosamine-N-acetylmuramyl-(pentapeptide)pyrophosphoryl-	|UDP-N-acetylglucosamine| + |MurAc(oyl-L-Ala-D-gamma-Glu-L-Lys-D-Ala-D-Ala)-diphospho- undecapre	2.4.1.227	R05662	0
UDP-N-acetylglucosamine-N-acetylmuramyl-(pentapeptide)pyrophosphoryl-	|UDP-N-acetylglucosamine| + |Undecaprenyl-diphospho-N-acetylmuramoyl-L-alanyl-D-glutamyl-meso-2	2.4.1.227	R05032	0.286573
D-AlanineD-Alanine ligase (ADP-forming)	|ATP| + (2) |D-Alanine| => |ADP| + |Phosphate| + |H+| + |Ala-Ala|	6.3.2.4	R01150	0.286573
UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-meso-2,6-	|ATP| + |Ala-Ala| + |UDP-N-acetylmuramoyl-L-alanyl-D-gamma-glutamyl-meso-2-6-diaminopimelate| =	6.3.2.10,6.3.2.15	R04617	0.286573
CoA:apo-[acyl-carrier-protein] pantetheinephosphotransferase	|CoA| + |apo-ACP| <=> |Adenosine 3-5-bisphosphate| + |ACP|	2.7.8.7	R01625	1.25981
L-Alanine racemase	|L-Alanine| <=> |D-Alanine|	5.1.1.1	R00401	-1000
Nitrous-oxide:(acceptor) oxidoreductase (NO-forming)	|H+| + |Quinol| + (2) |NO| <=> |H2O| + |Chinone| + |Nitrous oxide|	1.7.99.7	None	0
Xanthosine-5'-phosphate:L-glutamine amido-ligase (AMP-forming)	|H2O| + |ATP| + |L-Glutamine| + |XMP| => |PPi| + |AMP| + |L-Glutamate| + |H+| + |GMP|	6.3.5.2	R01231	12.6396
1-(5-Phospho-D-ribosyl)-5-amino-4-imidazolecarboxylate carboxy-lyase	|H+| + |5'-Phosphoribosyl-4-carboxy-5-aminoimidazole| => |CO2| + |AIR|	4.1.1.21	R04209	0
NCAIR synthetase and NCAIR mutase	|ATP| + |H2CO3| + |AIR| => |ADP| + |Phosphate| + |H+| + |5'-Phosphoribosyl-4-carboxy-5-amin	Undetermined	None	0
1-(5'-Phosphoribosyl)-5-amino-4-(N-succinocarboxamide)-imidazole	|SAICAR| <=> |Fumarate| + |AICAR|	4.3.2.2	R04559	0
N6-(1,2-Dicarboxyethyl)AMP AMP-lyase	|Adenylosuccinate| <=> |AMP| + |Fumarate|	4.3.2.2	R01083	15.1484
1-(5-Phosphoribosyl)-5-amino-4-carboxyimidazole:L-aspartate ligase	|ATP| + |L-Aspartate| + |5'-Phosphoribosyl-4-carboxy-5-aminoimidazole| => |ADP| + |Phosphate	6.3.2.6	R04591	0
5'-Phosphoribosylformylglycinamide:L-glutamine amido-ligase	|H2O| + |ATP| + |L-Glutamine| + |N-Formyl-GAR| => |ADP| + |Phosphate| + |L-Glutamate| + |H+| +	6.3.5.3	R04463	0
5-Phosphoribosylamine:pyrophosphate phosphoribosyltransferase	|PPi| + |L-Glutamate| + |5-Phosphoribosylamine| <= |H2O| + |L-Glutamine| + |PRPP|	2.4.2.14	R01072	0
2-(Formamido)-N1-(5-phosphoribosyl)acetamidine cyclo-ligase	|ATP| + |5'-Phosphoribosylformylglycinamidine| <=> |ADP| + |Phosphate| + |H+| + |AIR|	6.3.3.1	R04208	0
10-Formyltetrahydrofolate:5'-phosphoribosylglycinamide	|10-Formyltetrahydrofolate| + |GAR| <=> |H+| + |Tetrahydrofolate| + |N-Formyl-GAR|	2.1.2.2	R04325	0
IMP 1,2-hydrolase (decyclizing)	|H2O| + |IMP| <=> |FAICAR|	3.5.4.10	R01127	0
10-Formyltetrahydrofolate:5'-phosphoribosyl-5-amino-4-	|10-Formyltetrahydrofolate| + |AICAR| <=> |Tetrahydrofolate| + |FAICAR|	2.1.2.3	R04560	0
5-Phospho-D-ribosylamine:glycine ligase (ADP-forming)	|ATP| + |Glycine| + |5-Phosphoribosylamine| => |ADP| + |Phosphate| + |H+| + |GAR|	6.3.4.13	R04144	0
Adenine aminohydrolase	|H2O| + |H+| + |Adenine| => |NH3| + |HYXN|	3.5.4.2	R01244	0
L-Glutamate 5-semialdehyde:NAD+ oxidoreductase	|H2O| + |NAD| + |L-Glutamate5-semialdehyde| => |NADH| + |L-Glutamate| + (2) |H+|	1.5.1.12	R00245	0
L-1-Pyrroline-5-carboxylate:NAD+ oxidoreductase	(2) |H2O| + |NAD| + |1-Pyrroline-5-carboxylate| <=> |NADH| + |L-Glutamate| + |H+|	1.5.1.12	R00707	1000
2-methylpropionyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (dec	|H+| + |Malonyl-acyl-carrierprotein-| + |isobutyryl-ACP| => |CO2| + |ACP| + |4-methyl-3-oxo-pentano	2.3.1.0	None	0
4-methyl-pentanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (de	|H+| + |Malonyl-acyl-carrierprotein-| + |4-methyl-pentanoyl-ACP| => |CO2| + |ACP| + |6-methyl-3-oxc	2.3.1.0	None	0
6-methyl-heptanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (de	|H+| + |Malonyl-acyl-carrierprotein-| + |6-methyl-heptanoyl-ACP| => |CO2| + |ACP| + |8-methyl-3-oxc	2.3.1.0	None	0
8-methyl-nonanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (dec	|H+| + |Malonyl-acyl-carrierprotein-| + |8-methyl-nonanoyl-ACP| => |CO2| + |ACP| + |10-methyl-3-ox	2.3.1.0	None	0
10-methyl-undecanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (	|H+| + |Malonyl-acyl-carrierprotein-| + |10-methyl-undecanoyl-ACP| => |CO2| + |ACP| + |12-methyl-3	2.3.1.0	None	0
12-methyl-tridecanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (	|H+| + |Malonyl-acyl-carrierprotein-| + |12-methyl-tridecanoyl-ACP| => |CO2| + |ACP| + |14-methyl-3	2.3.1.0	None	0
Octanoyl-[acyl-carrier protein]:malonyl-[acyl-carrier-protein]	|H+| + |Octanoyl-ACP| + |Malonyl-acyl-carrierprotein-| => |CO2| + |3-oxodecanoyl-acp| + |ACP|	2.3.1.41,2.3.1.85,2.3.1.8	R04960	13.193
Tetradecanoyl-[acyl-carrier protein]:malonyl-[acyl-carrier-protein]	|H+| + |Myristoyl-ACP| + |Malonyl-acyl-carrierprotein-| => |CO2| + |3-oxohexadecanoyl-acp| + |ACP|	2.3.1.41,2.3.1.85,2.3.1.8	R04968	13.193
dodecanoyl-[acyl-carrier-protein]:malonyl-[acyl-carrier-protein]	|H+| + |Dodecanoyl-ACP| + |Malonyl-acyl-carrierprotein-| => |CO2| + |3-oxotetradecanoyl-acp| + |ACI	2.3.1.41,2.3.1.85,2.3.1.8	R04726	13.193
butyryl-[acyl-carrier protein]:malonyl-[acyl-carrier-protein]	|H+| + |Butyryl-ACP| + |Malonyl-acyl-carrierprotein-| => |CO2| + |3-Oxohexanoyl-[acp]| + |ACP|	2.3.1.41,2.3.1.85,2.3.1.8	R04952	13.193
Acyl-[acyl-carrier protein]:malonyl-[acyl-carrier-protein]	|H+| + |Malonyl-acyl-carrierprotein-| + |Acetyl-ACP| => |CO2| + |Acetoacetyl-ACP| + |ACP|	2.3.1.41,2.3.1.85,2.3.1.8	R04355	13.193
Decanoyl-[acyl-carrier protein]:malonyl-[acyl-carrier-protein]	|H+| + |Decanoyl-ACP| + |Malonyl-acyl-carrierprotein-| => |CO2| + |3-oxododecanoyl-acp| + |ACP|	2.3.1.41,2.3.1.85,2.3.1.8	R04963	13.193
acetyl-CoA:[acyl-carrier-protein] S-acetyltransferase	|Acetyl-CoA| + |ACP| <=> |CoA| + |Acetyl-ACP|	2.3.1.179,2.3.1.180,2.3.1	R01624	13.193
hexanoyl-[acyl-carrier protein]:malonyl-[acyl-carrier-protein]	|H+| + |Hexanoyl-ACP| + |Malonyl-acyl-carrierprotein-| => |CO2| + |3-oxooctanoyl-acp| + |ACP|	2.3.1.41,2.3.1.85,2.3.1.8	R04957	13.193
4-methyl-3-oxo-hexanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransfera	|H+| + |Malonyl-acyl-carrierprotein-| + |2-methylbutyryl-ACP| => |CO2| + |ACP| + |4-methyl-3-oxo-he	2.3.1.0	None	13.193
4-methyl-hexanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (dec	|H+| + |Malonyl-acyl-carrierprotein-| + |4-methyl-hexanoyl-ACP| => |CO2| + |ACP| + |6-methyl-3-oxo-	2.3.1.0	None	13.193
6-methyl-octanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (dec	|H+| + |Malonyl-acyl-carrierprotein-| + |6-methyl-octanoyl-ACP| => |CO2| + |ACP| + |8-methyl-3-oxo-	2.3.1.0	None	13.193
8-methyl-decanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (dec	|H+| + |Malonyl-acyl-carrierprotein-| + |8-methyl-decanoyl-ACP| => |CO2| + |ACP| + |10-methyl-3-oxc	2.3.1.0	None	13.193
10-methyl-dodecanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (	|H+| + |Malonyl-acyl-carrierprotein-| + |10-methyl-dodecanoyl-ACP| => |CO2| + |ACP| + |12-methyl-3	2.3.1.0	None	13.193
12-methyl-tetra-decanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransfera	|H+| + |Malonyl-acyl-carrierprotein-| + |12-methyl-tetra-decanoyl-ACP| => |CO2| + |ACP| + |14-meth	2.3.1.0	None	13.193
3-methylbutanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (decz	|H+| + |Malonyl-acyl-carrierprotein-| + |isovaleryl-ACP| => |CO2| + |ACP| + |5-methyl-3-oxo-hexanoyl	2.3.1.0	None	13.193
5-methyl-hexanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (dec	|H+| + |Malonyl-acyl-carrierprotein-| + |5-methyl-hexanoyl-ACP| => |CO2| + |ACP| + |7-methyl-3-oxo-	2.3.1.0	None	13.193
7-methyl-octanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (dec	|H+| + |Malonyl-acyl-carrierprotein-| + |7-methyl-octanoyl-ACP| => |CO2| + |ACP| + |9-methyl-3-oxo-	2.3.1.0	None	13.193
9-methyl-decanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (dec	|H+| + |Malonyl-acyl-carrierprotein-| + |9-methyl-decanoyl-ACP| => |CO2| + |ACP| + |11-methyl-3-oxc	2.3.1.0	None	13.193
11-methyl-dodecanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (	|H+| + |Malonyl-acyl-carrierprotein-| + |11-methyl-dodecanoyl-ACP| => |CO2| + |ACP| + |13-methyl-3	2.3.1.0	None	13.193
13-methyl-tetra-decanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransfera	|H+| + |Malonyl-acyl-carrierprotein-| + |13-methyl-tetra-decanoyl-ACP| => |CO2| + |ACP| + |15-meth	2.3.1.0	None	13.193
3-Oxooctodecanoyl-ACP:malonyl-[acyl-carrier-protein] C-acyltransferase (de	|H+| + |hexadecanoyl-acp| + |Malonyl-acyl-carrierprotein-| => |CO2| + |ACP| + |3-Oxooctodecanoyl-A	2.3.1.0	None	13.193
nitrate transport in via proton symport	|H+[e]| + |Nitrate[e]| <=> |H+| + |Nitrate|	Undetermined	None	-997.48
nitrite transport in via proton symport	|H+[e]| + |Nitrite[e]| <=> |H+| + |Nitrite|	Undetermined	None	997.48
Reduced azurin:oxygen oxidoreductase	|H2O| + |NO| + |Oxidized azurin| <=> (2) |H+| + |Nitrite| + |Reduced azurin|	1.7.2.1	R00785	0
D-Methionine-ABC transport	|H2O| + |ATP| + |D-Methionine[e]| => |ADP| + |Phosphate| + |H+| + |D-Methionine|	Undetermined	None	0
L-Methionine ABC transport	|H2O| + |ATP| + |L-Methionine[e]| => |ADP| + |Phosphate| + |L-Methionine| + |H+|	Undetermined	None	0
L-methionine S-oxide transport via ABC system	|H2O| + |ATP| + |L-Methionine S-oxide[e]| => |ADP| + |Phosphate| + |H+| + |L-Methionine S-oxide|	Undetermined	None	0
L-methionine R-oxide transport via ABC system	|H2O| + |ATP| + |L-methionine R-oxide[e]| => |ADP| + |Phosphate| + |H+| + |L-methionine R-oxide|	Undetermined	None	0
ATP:L-rhamnulose 1-phosphotransferase	|ATP| + |L-Lyxulose| <=> |ADP| + |L-Xylulose 1-phosphate|	2.7.1.5	R01902	0
ATP:L-rhamnulose 1-phosphotransferase	|ATP| + |L-Rhamnulose| <=> |ADP| + |L-Rhamnulose 1-phosphate|	2.7.1.5	R03014	0
L-Rhamnose ketol-isomerase	|L-Rhamnose| <=> |L-Rhamnulose|	5.3.1.14	R02437	0
L-Rhamnulose-1-phosphate lactaldehyde-lyase	|L-Rhamnulose 1-phosphate| <=> |Glycerone-phosphate| + |L-Lactaldehyde|	4.1.2.19	R02263	0
Putrescine-ABC transport	|H2O| + |ATP| + |Putrescine[e]| => |ADP| + |Phosphate| + |H+| + |Putrescine|	Undetermined	None	1.25981
Spermidine-ABC transport	|H2O| + |ATP| + |Spermidine[e]| => |ADP| + |Phosphate| + |H+| + |Spermidine|	Undetermined	None	1.25981
Glucose-phosphotransferase (PTS) system	|D-Glucose[e]| + |Phosphoenolpyruvate| <=> |Pyruvate| + |D-glucose-6-phosphate|	Undetermined	None	770.125
maltose transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |Maltose[e]| <=> |Pyruvate| + |maltose-6-phosphate|	Undetermined	None	0
Maltose-6'-phosphate 6-phosphoglucohydrolase	|H2O| + |maltose-6-phosphate| => |D-Glucose| + |D-glucose-6-phosphate|	3.2.1.122	R00838	0
(S)-Malate:(acceptor) oxidoreductase	|FAD| + |L-Malate| <=> |Oxaloacetate| + |FADH2|	1.1.99.16	R01257	-965.611
Malate dehydrogenase (ubiquinone 8 as acceptor)	|L-Malate| + |Ubiquinone-8| => |Oxaloacetate| + |Ubiquinol-8|	1.1.99.16	None	0
Malate dehydrogenase (menaquinone 8 as acceptor)	|L-Malate| + |Menaquinone 8| => |Oxaloacetate| + |Menaquinol 8|	1.1.99.16	None	0
4-amino-5-hydroxymethyl-2-methylpyrimidine synthetase	|H2O| + |AIR| <=> (0.5) |O2| + |Phosphate| + (2) |H+| + |Glycolaldehyde| + |Toxopyrimidine|	Undetermined	None	0
Succinate-semialdehyde:NADP+ oxidoreductase	|H2O| + |NADP| + |4-Oxobutanoate| => |NADPH| + |Succinate| + (2) |H+|	1.2.1.16	R00714	0
(S)-lactaldehyde:NAD+ oxidoreductase	|H2O| + |NAD| + |L-Lactaldehyde| <=> |NADH| + (2) |H+| + |L-Lactate|	1.2.1.21,1.2.1.22	R01446	0
N-Acetyl-L-glutamate-5-semialdehyde:NADP+ 5-oxidoreductase	|NADP| + |Phosphate| + |2-Acetamido-5-oxopentanoate| <= |NADPH| + |n-acetylglutamyl-phosphate|	1.2.1.38	R03443	-30.687
Xanthine:NAD+ oxidoreductase	|H2O| + |NAD| + |HYXN| <=> |NADH| + |H+| + |XAN|	1.17.1.4,1.1.1.204	R01768	-1000
xanthine:NAD+ oxidoreductase	|H2O| + |NAD| + |XAN| => |NADH| + |H+| + |Urate|	1.17.1.4,1.1.1.204	R02103	0
Acetyl-CoA:L-glutamate N-acetyltransferase	|Acetyl-CoA| + |L-Glutamate| => |CoA| + |H+| + |N-Acetyl-L-glutamate|	2.3.1.1	R00259	0
N2-Acetyl-L-ornithine:L-glutamate N-acetyltransferase	|L-Glutamate| + |N-Acetylornithine| <=> |Ornithine| + |N-Acetyl-L-glutamate|	2.3.1.35	R02282	30.687
ATP:N-acetyl-L-glutamate 5-phosphotransferase	|ATP| + |H+| + |N-Acetyl-L-glutamate| <=> |ADP| + |n-acetylglutamyl-phosphate|	2.7.2.8	R02649	30.687
N2-Acetyl-L-ornithine:2-oxoglutarate aminotransferase	|2-Oxoglutarate| + |N-Acetylornithine| <=> |L-Glutamate| + |2-Acetamido-5-oxopentanoate|	2.6.1.11	R02283	-30.687
Acetyl phosphate phosphohydrolase	|H2O| + |Acetylphosphate| => |Phosphate| + |Acetate| + (2) |H+|	3.6.1.7	R00317	0
3-Phospho-D-glyceroyl phosphate phosphohydrolase	|H2O| + |1,3-Bisphospho-D-glycerate| => |Phosphate| + (2) |H+| + |3-Phosphoglycerate|	3.6.1.7	R01515	0
S-Adenosyl-L-methionine:uroporphyrin-III C-methyltransferase	(2) |H+| + |Siroheme| <=> |Sirohydrochlorin| + |Fe2+|	4.99.1.4	R02864	-1.25981
S-Adenosyl-L-methionine:uroporphyrin-III C-methyltransferase	(2) |S-Adenosyl-L-methionine| + |UroporphyrinogenIII| <=> (2) |S-Adenosyl-homocysteine| + (2) |H+| +	2.1.1.107	R03194	1.25981
S-Adenosyl-L-methionine:uroporphyrin-III C-methyltransferase	|NAD| + |Precorrin 2| <=> |NADH| + |H+| + |Sirohydrochlorin|	1.3.1.76,2.1.1.107	R03947	1.25981
Hydrogen-sulfide:ferredoxin oxidoreductase	(3) |H2O| + |H2S| + (3) |Oxidizedferredoxin| <=> (7) |H+| + |Sulfite| + (3) |Reducedferredoxin|	1.8.7.1	None	0
ATP:adenylylsulfate 3'-phosphotransferase	|ATP| + |APS| <=> |ADP| + |3-phosphoadenylylsulfate|	2.7.1.25	R00509	-1.25981
ATP:sulfate adenylyltransferase	|ATP| + |Sulfate| + |H+| <=> |PPi| + |APS|	2.7.7.4	R00529	-1.25981
adenosine 3',5'-bisphosphate,sulfite:oxidized-thioredoxin oxidore	|3-phosphoadenylylsulfate| + |trdrd| <=> |Adenosine 3-5-bisphosphate| + |Sulfite| + |trdox|	1.8.4.8	R02021	-1.25981
APSPTi	|APS| + |trdrd| => |AMP| + |Sulfite| + |trdox|	1.8.4.10	R07176	0
Carbon-dioxide:L-glutamine amido-ligase (ADP-forming,	|H2O| + (2) |ATP| + |L-Glutamine| + |H2CO3| => (2) |ADP| + |Phosphate| + |L-Glutamate| + |H+| + |(	6.3.5.5	R00575	0
calcium transport in/out via proton antiporter	|Ca2+| + |H+[e]| <=> |Ca2+[e]| + |H+|	TC-2.A.19,2.A.19	None	-1.25981
(S)-Malate hydro-lyase	|L-Malate| <=> |H2O| + |Fumarate|	4.2.1.2	R01082	709.337
Carbamoyl-phosphate:L-ornithine carbamoyltransferase	|Ornithine| + |Carbamoylphosphate| => |Phosphate| + (2) |H+| + |Citrulline|	2.1.3.3	R01398	0
Butanoyl-CoA:(acceptor) 2,3-oxidoreductase	|NAD| + |Butyryl-CoA| <= |NADH| + |H+| + |Crotonyl-CoA|	1.3.1.44,1.3.99.2	R01171	0

Reactions with corresponding equation, EC number and FBA flux value under aerobic conditions in rich media

2-Methylpropanoyl-CoA:oxygen 2,3-oxidoreductase	|O2| + (2) |Isobutyryl-CoA| => (2) |H2O| + (2) |Methacrylyl-CoA|	1.3.99.2	None	0
(S)-2-methylbutanoyl-CoA:oxygen 2,3-oxidoreductase	|O2| + (2) |2-Methylbutyryl-CoA| => (2) |H2O| + (2) |Tiglyl-CoA|	1.3.99.2	None	0
Dipeptide transport via ABC system (ala-asp)	|H2O| + |ATP| + |ala-L-asp-L[e]| => |ADP| + |Phosphate| + |H+| + |ala-L-asp-L|	TC-3.A.1.5,3.A.1.5	None	0
Dipeptide transport via ABC system (gly-glu)	|H2O| + |ATP| + |gly-glu-L[e]| => |ADP| + |Phosphate| + |H+| + |gly-glu-L|	TC-3.A.1.5,3.A.1.5	None	0
Dipeptide transport via ABC system (met-ala)	|H2O| + |ATP| + |met-L-ala-L[e]| => |ADP| + |Phosphate| + |H+| + |met-L-ala-L|	TC-3.A.1.5,3.A.1.5	None	0
Dipeptide transport via ABC system (gly-asp)	|H2O| + |ATP| + |gly-asp-L[e]| => |ADP| + |Phosphate| + |H+| + |gly-asp-L|	TC-3.A.1.5,3.A.1.5	None	0
Dipeptide transport via ABC system (gly-pro-L)	|H2O| + |ATP| + |gly-pro-L[e]| => |ADP| + |Phosphate| + |H+| + |gly-pro-L|	TC-3.A.1.5,3.A.1.5	None	0
Dipeptide transport via ABC system (cgly)	|H2O| + |ATP| + |Cys-Gly[e]| => |ADP| + |Phosphate| + |H+| + |Cys-Gly|	TC-3.A.1.5,3.A.1.5	None	0
Dipeptide transport via ABC system (ala-gln)	|H2O| + |ATP| + |Ala-Gln[e]| => |ADP| + |Phosphate| + |H+| + |Ala-Gln|	TC-3.A.1.5,3.A.1.5	None	0
Dipeptide transport via ABC system (ala-glu)	|H2O| + |ATP| + |ala-L-glu-L[e]| => |ADP| + |Phosphate| + |H+| + |ala-L-glu-L|	TC-3.A.1.5,3.A.1.5	None	0
Dipeptide transport via ABC system (ala-gly)	|H2O| + |ATP| + |L-alanylglycine[e]| => |ADP| + |Phosphate| + |H+| + |L-alanylglycine|	TC-3.A.1.5,3.A.1.5	None	0
Dipeptide transport via ABC system (ala-leu)	|H2O| + |ATP| + |Ala-Leu[e]| => |ADP| + |Phosphate| + |H+| + |Ala-Leu|	TC-3.A.1.5,3.A.1.5	None	0
Dipeptide transport via ABC system (gly-gln)	|H2O| + |ATP| + |Gly-Gln[e]| => |ADP| + |Phosphate| + |H+| + |Gly-Gln|	TC-3.A.1.5,3.A.1.5	None	0
Gly-Phe ABC transporters	|H2O| + |ATP| + |Gly-Phe[e]| => |ADP| + |Phosphate| + |Gly-Phe|	3.A.1.5	None	0
Gly-Try ABC transporters	|H2O| + |ATP| + |Gly-Tyr[e]| => |ADP| + |Phosphate| + |Gly-Tyr|	3.A.1.5	None	0
Gly-Cys ABC transporters	|H2O| + |ATP| + |Gly-Cys[e]| => |ADP| + |Phosphate| + |Gly-Cys|	3.A.1.5	None	12.8165
Dipeptide transport via ABC system (ala-his)	|H2O| + |ATP| + |Ala-His[e]| => |ADP| + |Phosphate| + |H+| + |Ala-His|	TC-3.A.1.5,3.A.1.5	None	13.003
Dipeptide transport via ABC system (gly-met)	|H2O| + |ATP| + |Gly-Met[e]| => |ADP| + |Phosphate| + |H+| + |Gly-Met|	TC-3.A.1.5,3.A.1.5	None	19.2689
Dipeptide transport via ABC system (gly-asn)	|H2O| + |ATP| + |gly-asn-L[e]| => |ADP| + |Phosphate| + |H+| + |gly-asn-L|	TC-3.A.1.5,3.A.1.5	None	23.5354
Dipeptide transport via ABC system (ala-thr)	|H2O| + |ATP| + |ala-L-Thr-L[e]| => |ADP| + |Phosphate| + |H+| + |ala-L-Thr-L|	TC-3.A.1.5,3.A.1.5	None	29.6468
Gly-Leu ABC transporters	|H2O| + |ATP| + |Gly-Leu[e]| => |ADP| + |Phosphate| + |Gly-Leu|	3.A.1.5	None	47.8096
formate transport in via proton symport	|Formate[e]| + |H+[e]| <=> |Formate| + |H+|	Undetermined	None	-950.569
Tetradecanoyl-[acyl-carrier protein]:malonyl-CoA	|NADP| + |Myristoyl-ACP| <= |NADPH| + |H+| + |(2E)-Tetradecenoyl-[acp]|	1.3.1.10,2.3.1.85	R04967	0
Octanoyl-[acyl-carrier protein]:malonyl-CoA	|NADP| + |Octanoyl-ACP| <= |NADPH| + |H+| + |(2E)-Octenoyl-[acp]|	1.3.1.10,2.3.1.85	R04959	0
Butyryl-[acyl-carrier protein]:malonyl-CoA	|NADP| + |Butyryl-ACP| <= |NADPH| + |H+| + |But-2-enoyl-[acyl-carrier protein]|	1.3.1.10	R04430	0
Hexadecanoyl-[acyl-carrier protein:malonyl-CoA	|NADP| + |hexadecanoyl-acp| <= |NADPH| + |H+| + |(2E)-Hexadecenoyl-[acp]|	1.3.1.10,2.3.1.85	R04970	0
Dodecanoyl-[acyl-carrier protein]: malonyl-CoA	|NADP| + |Dodecanoyl-ACP| <= |NADPH| + |H+| + |(2E)-Dodecenoyl-[acp]|	1.3.1.10,2.3.1.85	R04725	0
Decanoyl-[acyl-carrier protein]:malonyl-CoA	|NADP| + |Decanoyl-ACP| <= |NADPH| + |H+| + |(2E)-Decenoyl-[acp]|	1.3.1.10,2.3.1.85	R04962	0
Hexanoyl-[acyl-carrier protein]:oxoacyl- and enoyl-reducing and	|NADP| + |Hexanoyl-ACP| <= |NADPH| + |H+| + |(2E)-Hexenoyl-[acp]|	1.3.1.10,2.3.1.85	R04956	0
4-methyl-trans-hex-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |4-methyl-trans-hex-2-enoyl-ACP| => |NADP| + |4-methyl-hexanoyl-ACP|	1.3.1.0	None	0
6-methyl-trans-oct-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |6-methyl-trans-oct-2-enoyl-ACP| => |NADP| + |6-methyl-octanoyl-ACP|	1.3.1.0	None	0
8-methyl-trans-dec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |8-methyl-trans-dec-2-enoyl-ACP| => |NADP| + |8-methyl-decanoyl-ACP|	1.3.1.0	None	0
10-methyl-trans-dodec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |10-methyl-trans-dodec-2-enoyl-ACP| => |NADP| + |10-methyl-dodecanoyl-ACP|	1.3.1.0	None	0
12-methyl-trans-tetra-dec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |12-methyl-trans-tetra-dec-2-enoyl-ACP| => |NADP| + |12-methyl-tetra-decanoyl-A(	1.3.1.0	None	0
14-methyl-trans-hexa-dec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |14-methyl-trans-hexa-dec-2-enoyl-ACP| => |NADP| + |14-methyl-hexa-decanoyl-AC	1.3.1.0	None	0
5-methyl-trans-hex-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |5-methyl-trans-hex-2-enoyl-ACP| => |NADP| + |5-methyl-hexanoyl-ACP|	1.3.1.0	None	0
7-methyl-trans-oct-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |7-methyl-trans-oct-2-enoyl-ACP| => |NADP| + |7-methyl-octanoyl-ACP|	1.3.1.0	None	0
9-methyl-trans-dec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |9-methyl-trans-dec-2-enoyl-ACP| => |NADP| + |9-methyl-decanoyl-ACP|	1.3.1.0	None	0
11-methyl-trans-dodec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |11-methyl-trans-dodec-2-enoyl-ACP| => |NADP| + |11-methyl-dodecanoyl-ACP|	1.3.1.0	None	0
13-methyl-trans-tetra-dec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |13-methyl-trans-tetra-dec-2-enoyl-ACP| => |NADP| + |13-methyl-tetra-decanoyl-A(	1.3.1.0	None	0
15-methyl-trans-hexa-dec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |15-methyl-trans-hexa-dec-2-enoyl-ACP| => |NADP| + |15-methyl-hexa-decanoyl-AC	1.3.1.0	None	0
4-methyl-trans-pent-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |4-methyl-trans-pent-2-enoyl-ACP| => |NADP| + |4-methyl-pentanoyl-ACP|	1.3.1.0	None	0
6-methyl-trans-hept-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |6-methyl-trans-hept-2-enoyl-ACP| => |NADP| + |6-methyl-heptanoyl-ACP|	1.3.1.0	None	0
8-methyl-trans-non-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |8-methyl-trans-non-2-enoyl-ACP| => |NADP| + |8-methyl-nonanoyl-ACP|	1.3.1.0	None	0
10-methyl-trans-undec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |10-methyl-trans-undec-2-enoyl-ACP| => |NADP| + |10-methyl-undecanoyl-ACP|	1.3.1.0	None	0
12-methyl-trans-tridec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |12-methyl-trans-tridec-2-enoyl-ACP| => |NADP| + |12-methyl-tridecanoyl-ACP|	1.3.1.0	None	0
14-methyl-trans-pentadec-2-enoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |14-methyl-trans-pentadec-2-enoyl-ACP| => |NADP| + |14-methyl-pentadecanoyl-A	1.3.1.0	None	0
Octodecanoyl-ACP:NADP+ oxidoreductase (A-specific)	|NADPH| + |H+| + |trans-Octodec-2-enoyl-ACP| => |NADP| + |Octodecanoyl-ACP|	1.3.1.0	None	0
L-Glutamate:NADP+ oxidoreductase (transaminating)	|NADP| + (2) |L-Glutamate| <=> |NADPH| + |2-Oxoglutarate| + |L-Glutamine| + |H+|	1.4.1.13	R00114	724.938
(S)-Lactate:NAD+ oxidoreductase	|NAD| + |L-Lactate| <=> |NADH| + |Pyruvate| + |H+|	1.1.1.27	R00703	-1000
L-Lactate dehydrogenase (ubiquinone)	|L-Lactate| + |Ubiquinone-8| => |Pyruvate| + |Ubiquinol-8|	1.1.2.3	None	0
L-Lactate dehydrogenase (menaquinone)	|L-Lactate| + |Menaquinone 8| => |Pyruvate| + |Menaquinol 8|	1.1.2.3	None	0
(S)-Lactate:ferricytochrome-c 2-oxidoreductase	(2) |Cytochrome c3+| + |L-Lactate| <=> |Pyruvate| + (2) |H+| + (2) |Cytochrome c2+|	1.1.2.3	R00196	188.625
L-lactate reversible transport via proton symport	|H+[e]| + |L-Lactate[e]| <=> |H+| + |L-Lactate|	Undetermined	None	-811.375
palmitoyl-cardiolipin synthase	(2) |Phosphatidylglycerol dihexadecanoyl| <=> |Glycerol| + |Palmitoylcardiolipin (B. subtilis)|	Undetermined	None	0
myristoyl-cardiolipin synthase	(2) |Phosphatidylglycerol ditetradecanoyl| <=> |Glycerol| + |Myristoylcardiolipin (B. subtilis)|	Undetermined	None	0
isotetradecanoyl-cardiolipin synthase	(2) |Diisotetradecanoylphosphatidylglycerol| <=> |Glycerol| + |Isotetradecanoylcardiolipin (B. subtilis)|	Undetermined	None	0
isopentadecanoyl-cardiolipin synthase	(2) |Diisopentadecanoylphosphatidylglycerol| <=> |Glycerol| + |Isopentadecanoylcardiolipin (B. subtilis	Undetermined	None	0
anteisopentadecanoyl-cardiolipin synthase	(2) |Dianteisopentadecanoylphosphatidylglycerol| <=> |Glycerol| + |Anteisopentadecanoylcardiolipin (I	Undetermined	None	0
isohexadecanoyl-cardiolipin synthase	(2) |Diisohexadecanoylphosphatidylglycerol| <=> |Glycerol| + |Isohexadecanoylcardiolipin (B. subtilis)|	Undetermined	None	0
stearoyl-cardiolipin synthase	(2) |Phosphatidylglycerol dioctadecanoyl| <=> |Glycerol| + |Stearoylcardiolipin (B. subtilis)|	Undetermined	None	1.36255
isoheptadecanoyl-cardiolipin synthase	(2) |Diisoheptadecanoylphosphatidylglycerol| <=> |Glycerol| + |Isoheptadecanoylcardiolipin (B. subtilis	Undetermined	None	1.36255
anteisoheptadecanoyl-cardiolipin synthase	(2) |Dianteisoheptadecanoylphosphatidylglycerol| <=> |Glycerol| + |Anteisoheptadecanoylcardiolipin (I	Undetermined	None	1.36255
ATP pyrophosphate-lyase (cyclizing)	|ATP| => |PPi| + |cAMP|	4.6.1.1	R00089	0
Butyryl-[acyl-carrier protein]:malonyl-CoA	|NAD| + |Butyryl-ACP| <= |NADH| + |H+| + |But-2-enoyl-[acyl-carrier protein]|	1.3.1.9,2.3.1.86	R04429	-13.193
Tetradecanoyl-[acyl-carrier protein]:malonyl-CoA	|NAD| + |Myristoyl-ACP| <= |NADH| + |H+| + |(2E)-Tetradecenoyl-[acp]|	1.3.1.9,2.3.1.86	R04966	-13.193
Dodecanoyl-[acyl-carrier protein]:malonyl-CoA	|NAD| + |Dodecanoyl-ACP| <= |NADH| + |H+| + |(2E)-Dodecenoyl-[acp]|	1.3.1.9,2.3.1.86	R04724	-13.193
Octanoyl-[acyl-carrier protein]:malonyl-CoA	|NAD| + |Octanoyl-ACP| <= |NADH| + |H+| + |(2E)-Octenoyl-[acp]|	1.3.1.9,2.3.1.86	R04958	-13.193
Hexanoyl-[acyl-carrier protein]:malonyl-CoA	|NAD| + |Hexanoyl-ACP| <= |NADH| + |H+| + |(2E)-Hexenoyl-[acp]|	1.3.1.9,2.3.1.86	R04955	-13.193
Decanoyl-[acyl-carrier protein]:malonyl-CoA	|NAD| + |Decanoyl-ACP| <= |NADH| + |H+| + |(2E)-Decenoyl-[acp]|	1.3.1.9,2.3.1.86	R04961	-13.193
Hexadecanoyl-[acyl-carrier protein]:malonyl-CoA	|NAD| + |hexadecanoyl-acp| <= |NADH| + |H+| + |(2E)-Hexadecenoyl-[acp]|	1.3.1.9,2.3.1.86	R04969	-13.193
4-methyl-trans-pent-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |4-methyl-trans-pent-2-enoyl-ACP| => |NAD| + |4-methyl-pentanoyl-ACP|	1.3.1.0	None	0
6-methyl-trans-hept-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |6-methyl-trans-hept-2-enoyl-ACP| => |NAD| + |6-methyl-heptanoyl-ACP|	1.3.1.0	None	0
8-methyl-trans-non-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |8-methyl-trans-non-2-enoyl-ACP| => |NAD| + |8-methyl-nonanoyl-ACP|	1.3.1.0	None	0
10-methyl-trans-undec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |10-methyl-trans-undec-2-enoyl-ACP| => |NAD| + |10-methyl-undecanoyl-ACP|	1.3.1.0	None	0
12-methyl-trans-tridec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |12-methyl-trans-tridec-2-enoyl-ACP| => |NAD| + |12-methyl-tridecanoyl-ACP|	1.3.1.0	None	0
14-methyl-trans-pentadec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |14-methyl-trans-pentadec-2-enoyl-ACP| => |NAD| + |14-methyl-pentadecanoyl-ACP	1.3.1.0	None	0
4-methyl-trans-hex-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |4-methyl-trans-hex-2-enoyl-ACP| => |NAD| + |4-methyl-hexanoyl-ACP|	1.3.1.0	None	13.193
6-methyl-trans-oct-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |6-methyl-trans-oct-2-enoyl-ACP| => |NAD| + |6-methyl-octanoyl-ACP|	1.3.1.0	None	13.193
8-methyl-trans-dec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |8-methyl-trans-dec-2-enoyl-ACP| => |NAD| + |8-methyl-decanoyl-ACP|	1.3.1.0	None	13.193
10-methyl-trans-dodec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |10-methyl-trans-dodec-2-enoyl-ACP| => |NAD| + |10-methyl-dodecanoyl-ACP|	1.3.1.0	None	13.193
12-methyl-trans-tetra-dec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |12-methyl-trans-tetra-dec-2-enoyl-ACP| => |NAD| + |12-methyl-tetra-decanoyl-ACP	1.3.1.0	None	13.193
14-methyl-trans-hexa-dec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |14-methyl-trans-hexa-dec-2-enoyl-ACP| => |NAD| + |14-methyl-hexa-decanoyl-ACP|	1.3.1.0	None	13.193
5-methyl-trans-hex-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |5-methyl-trans-hex-2-enoyl-ACP| => |NAD| + |5-methyl-hexanoyl-ACP|	1.3.1.0	None	13.193
7-methyl-trans-oct-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |7-methyl-trans-oct-2-enoyl-ACP| => |NAD| + |7-methyl-octanoyl-ACP|	1.3.1.0	None	13.193
9-methyl-trans-dec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |9-methyl-trans-dec-2-enoyl-ACP| => |NAD| + |9-methyl-decanoyl-ACP|	1.3.1.0	None	13.193
11-methyl-trans-dodec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |11-methyl-trans-dodec-2-enoyl-ACP| => |NAD| + |11-methyl-dodecanoyl-ACP|	1.3.1.0	None	13.193
13-methyl-trans-tetra-dec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |13-methyl-trans-tetra-dec-2-enoyl-ACP| => |NAD| + |13-methyl-tetra-decanoyl-ACP	1.3.1.0	None	13.193
15-methyl-trans-hexa-dec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |15-methyl-trans-hexa-dec-2-enoyl-ACP| => |NAD| + |15-methyl-hexa-decanoyl-ACP|	1.3.1.0	None	13.193
trans-Octodec-2-enoyl-ACP:NAD+ oxidoreductase (A-specific)	|NADH| + |H+| + |trans-Octodec-2-enoyl-ACP| => |NAD| + |Octodecanoyl-ACP|	1.3.1.0	None	13.193
allantoin transport in via proton symport	|H+[e]| + |Allantoin[e]| <=> |H+| + |Allantoin|	Undetermined	None	0
Thiamine transport in via proton symport	|H+[e]| + |Thiamin[e]| <=> |H+| + |Thiamin|	Undetermined	None	1.25981
cytosine transport in via proton symport	|H+[e]| + |Cytosine[e]| <=> |H+| + |Cytosine|	Undetermined	None	829.041
O-Succinyl-L-homoserine succinate-lyase (adding cysteine)	|H2O| + |O-Succinyl-L-homoserine| => |NH3| + |Succinate| + |H+| + |2-Oxobutyrate|	2.5.1.48,4.2.99.9	R00999	0
O-Succinyl-L-homoserine succinate-lyase (adding cysteine)	|Succinate| + |Cystathionine| <=> |L-Cysteine| + |O-Succinyl-L-homoserine|	2.5.1.48,4.2.99.9	R02508,R03260	0
O-Acetyl-L-homoserine succinate-lyase (adding cysteine)	|L-Cysteine| + |O-Acetyl-L-homoserine| <=> |Acetate| + |Cystathionine|	2.5.1.49,4.2.99.9	R03217	0
Cystathionine L-homocysteine-lyase (deaminating)	|H2O| + |Cystathionine| => |NH3| + |Pyruvate| + |Homocysteine|	4.4.1.8	R01285,R01286	0
Acetaldehyde:NAD+ oxidoreductase	|H2O| + |NAD| + |Acetaldehyde| => |NADH| + |Acetate| + (2) |H+|	1.2.1.3,1.2.1.5	R00710	0
D-Glyceraldehyde:NAD+ oxidoreductase	|H2O| + |NAD| + |D-Glyceraldehyde| => |NADH| + (2) |H+| + |Glycerate|	1.2.1.3	R01752	0
4-Aminobutyraldehyde:NAD+ oxidoreductase	|H2O| + |NADP| + |4-Aminobutanal| => |NADPH| + (2) |H+| + |GABA|	1.2.1.3	R01986	0
4-aminobutanal:NAD+ 1-oxidoreductase	|H2O| + |NAD| + |4-Aminobutanal| => |NADH| + (2) |H+| + |GABA|	1.2.1.19,1.2.1.3	R02549	0
Imidazole acetaldehyde:NAD+ oxidoreductase	|H2O| + |NAD| + |Imidazole acetaldehyde| => |NADH| + (2) |H+| + |4-Imidazoleacetate|	1.2.1.3	R04065	0
aldehyde dehydrogenase (glyoxylate, NAD)	|H2O| + |NAD| + |Glyoxalate| <=> |NADH| + (2) |H+| + |Oxalate|	Undetermined	None	0
aldehyde dehydrogenase (aminoacetaldehyde, NAD)	|H2O| + |NAD| + |Aminoacetaldehyde| => |NADH| + |Glycine| + (2) |H+|	Undetermined	None	1.25981
O-succinylbenzoate-CoA synthase	|H2O| + |Succinylbenzoate| <= |SHCHC|	4.2.1.-,4.2.1.113	R04031	-2.51962
alpha-D-Glucoside glucohydrolase	|H2O| + |Sucrose| => |D-Glucose| + |D-Fructose|	3.2.1.20,3.2.1.26,3.2.1.4	R00802,R00801	0
Isomaltose 6-alpha-D-glucanohydrolase	|H2O| + |Brachiose| => (2) |D-Glucose|	3.2.1.10	None	0
Maltodextrin glucosidase (maltotriose)	|H2O| + |Amylotriose| <=> |D-Glucose| + |Maltose|	Undetermined	None	0
Maltodextrin glucosidase (maltotetraose)	|H2O| + |Maltotetraose| <=> |D-Glucose| + |Amylotriose|	Undetermined	None	0
Maltodextrin glucosidase (maltopentaose)	|H2O| + |Maltopentaose| <=> |D-Glucose| + |Maltotetraose|	Undetermined	None	0
Maltodextrin glucosidase (maltohexaose)	|H2O| + |Maltohexaose| <=> |D-Glucose| + |Maltopentaose|	Undetermined	None	0
Maltodextrin glucosidase (maltoheptaose)	|H2O| + |Maltoheptaose| <=> |D-Glucose| + |Maltohexaose|	Undetermined	None	0
Maltodextrin glucosidase (dextrin)	(5) |H2O| + |Dextrin| => (6) |D-Glucose|	Undetermined	None	0
alpha-D-Glucoside glucohydrolase	|H2O| + |Maltose| => (2) |D-Glucose|	3.2.1.20	R00028	0
D-Ribose ABC transport	|H2O| + |ATP| + |D-Ribose[e]| => |ADP| + |Phosphate| + |H+| + |D-Ribose|	3.A.1.2	None	0
Deoxyribose transport via ABC system	|H2O| + |ATP| + |Thyminose[e]| => |ADP| + |Phosphate| + |H+| + |Thyminose|	Undetermined	None	0
D-ribose transport out via ABC system	|H2O| + |ATP| + |D-Ribose| => |ADP| + |Phosphate| + |H+| + |D-Ribose[e]|	TC-3.A.1.2,3.A.1.2	None	0
5-Methylthio-5-deoxy-D-ribose-1-phosphate ketol-isomerase	|methylthioribose-1-phosphate| <=> |methylthioribulose-1-phosphate|	5.3.1.23	R04420	0

ATP:S5-methyl-5-thio-D-ribose 1-phosphotransferase	|ATP| + |5-Methylthio-D-ribose| <=> |ADP| + |methylthioribose-1-phosphate|	2.7.1.100	R04143	0
Acetamide amidohydrolase	|H2O| + |Acetamide| => |NH3| + |Acetate|	3.5.1.4	R00321	0
2-phenylacetamide amidohydrolase	|H2O| + |H+| + |2-Phenylacetamide| <=> |NH3| + |PACT|	3.5.1.4	R02540	0
Indole-3-acetamide amidohydrolase	|H2O| + |Indole-3-acetamide| => |NH3| + |Indoleacetate|	3.5.1.4	R03096	0
4-Guanidinobutanamide amidohydrolase	|H2O| + |4-Guanidinobutanamide| => |NH3| + |4-Guanidinobutanoate|	3.5.1.4	R03180	0
AMID5	|NH3| + |Acrylate| <= |H2O| + |Acrylamide|	3.5.1.4	R05551	0
Acylamide aminohydrolase	|H2O| + |Benzamide| => |NH3| + |Benzoate|	3.5.1.4	R05590	0
2-keto-4-methylthiobutyrate transamination	|L-Glutamate| + |4-methylthio 2-oxobutyrate| <=> |2-Oxoglutarate| + |L-Methionine|	2.6.1.5,2.6.1.57	R07396	0
2,3-diketo-5-methylthiopentyl-1-phosphate enolase	|2,3-diketo5-methylthio-1-phosphopentane| <=> |H+| + |2-Hydroxy-3-keto-5-methylthiopentenyl-1-ph	3.1.3.77	R07393	0
2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phophatase	|H2O| + |2-Hydroxy-3-keto-5-methylthiopentenyl-1-phosphate| => |Phosphate| + |H+| + |1,2-dihydro	3.1.3.77	R07394	0
S-Methyl-5-thio-D-ribulose-1-phosphate hydro-lyase	|methylthioribulose-1-phosphate| => |H2O| + |2,3-diketo5-methylthio-1-phosphopentane|	4.2.1.109	R07392	0
1,2-dihydroxy-5-(methylthio)pent-1-en-3-one:oxygen oxidoreductase	|O2| + |1,2-dihydroxy-3-keto-5-methylthiopentene| => |Formate| + |H+| + |4-methylthio 2-oxobutyra	1.13.11.54	R07364	0
arginine/ornithine antiporter	|L-Arginine[e]| + |Ornithine| <=> |L-Arginine| + |Ornithine[e]|	Undetermined	None	1000
NADH:6,7-dihydropteridine oxidoreductase	|NADH| + |H+| + |Dihydrobiopterin| <=> |NAD| + |Tetrahydrobiopterin|	1.5.1.34	R01793	-1000
NADPH:6,7-dihydropteridine oxidoreductase	|NADPH| + |H+| + |Dihydrobiopterin| <=> |NADP| + |Tetrahydrobiopterin|	1.5.1.34	R01794	1000
Proton sodium antiport	|H+| + |Na+[e]| <=> |H+[e]| + |Na+|	Undetermined	None	-1000
mannitol transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |D-Mannitol[e]| <=> |Pyruvate| + |D-mannitol-1-phosphate|	Undetermined	None	1000
trehalose transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |TRHL[e]| <=> |Pyruvate| + |Trehalose 6-phosphate|	2.7.1.69	None	-950.996
N-Acetyl-D-glucosamine transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |N-Acetyl-D-glucosamine[e]| <=> |Pyruvate| + |N-Acetyl-D-glucosamine 6-ph	Undetermined	None	-1000
D-glucosamine transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |GLUM[e]| <=> |Pyruvate| + |D-Glucosamine phosphate|	Undetermined	None	365.679
arbutin transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |Ursin[e]| <=> |Pyruvate| + |Arbutin-6P|	Undetermined	None	0
salicin transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |Salicin[e]| <=> |Pyruvate| + |Salicin-6P|	Undetermined	None	0
D-mannose transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |D-Mannose[e]| <=> |Pyruvate| + |D-mannose-6-phosphate|	Undetermined	None	-1000
sucrose transport via PEP:Pyr PTS	|Phosphoenolpyruvate| + |Sucrose[e]| <=> |Pyruvate| + |6-Phosphosucrose|	Undetermined	None	0
2-Oxoglutarate:Thiamin diphosphate 2-oxidoreductase(decarboxylating	|2-Oxoglutarate| + |TPP| + |H+| => |CO2| + |3-Carboxy-1-hydroxypropyl-ThPP|	1.2.4.2	R00621	0
3-Carboxy-1-hydroxypropyl-ThPP:lipoamde	|Lipoamide| + |3-Carboxy-1-hydroxypropyl-ThPP| <=> |TPP| + |S-Succinyldihydrolipoamide|	1.2.4.2	R03316	0
succinyl-CoA:enzyme N6-(dihydrolipoyl)lysine S-succinyltransferase	|Succinyl-CoA| + |Dihydrolipoamide| <=> |CoA| + |S-Succinyldihydrolipoamide|	2.3.1.61	R02570	0
Acetyl-L,L-diaminopimelate aminotransferase	|2-Oxoglutarate| + |N-acetyl-LL-2,6-diaminopimelate| <=> |L-Glutamate| + |L-2-Acetamido-6-oxopime	2.6.1.-	R04467	-0.286573
beta-D-Glucose:NADP+ 1-oxoreductase	|NADP| + |beta-D-Glucose| <=> |NADPH| + |H+| + |Gluconolactone|	1.1.1.47	R01521	-1000
beta-D-Glucose:NAD+ 1-oxoreductase	|NAD| + |beta-D-Glucose| <=> |NADH| + |H+| + |Gluconolactone|	1.1.1.47	R01520	1000
Acetyl-CoA:L-2,3,4,5-tetrahydrodipicolinate N2-acetyltransferase	|H2O| + |Acetyl-CoA| + |tetrahydrodipicolinate| <=> |CoA| + |L-2-Acetamido-6-oxopimelate|	2.3.1.89	R04364	0.286573
N6-Acetyl-LL-2,6-diaminoheptanedioate amidohydrolase	|H2O| + |N-acetyl-LL-2,6-diaminopimelate| <=> |Acetate| + |LL-2,6-Diaminopimelate|	3.5.1.47	R02733	0.286573
potassium transport out via proton antiport	|H+[e]| + |K+| <=> |H+| + |K+[e]|	TC-2.A.37,2.A.37	None	-1000
pyruvate:thiamin diphosphate acetaldehydetransferase	|CO2| + |2-Hydroxyethyl-ThPP| <= |Pyruvate| + |TPP| + |H+|	1.2.4.1,2.2.1.6,4.1.1.1	R00014	0
2-(alpha-Hydroxyethyl)thiamine diphosphate:lipoamide	|Lipoamide| + |2-Hydroxyethyl-ThPP| <=> |TPP| + |S-Acetyldihydrolipoamide|	1.2.4.1	R03270	-27.6337
acetyl-CoA:enzyme N6-(dihydrolipoyl)lysine S-acetyltransferase	|Acetyl-CoA| + |Dihydrolipoamide| <=> |CoA| + |S-Acetyldihydrolipoamide|	2.3.1.12	R02569	27.6337
Dihydrolipoamide:NAD+ oxidoreductase	|NAD| + |Dihydrolipoamide| <=> |NADH| + |H+| + |Lipoamide|	1.8.1.4	R07618,R01698	-1.24778
L-Lysine carboxy-lyase	|L-Lysine| + |H+| => |CO2| + |Cadaverine|	4.1.1.18	R00462	0
L-Arginine carboxy-lyase	|L-Arginine| + |H+| => |CO2| + |Agmatine|	4.1.1.19	R00566	0
myo-Inositol 1-phosphate phosphahydrolase	|H2O| + |Inositol 1-phosphate| => |Phosphate| + |H+| + |L-Inositol|	3.1.3.25	R01185	0
myo-Inositol 4-phosphate phosphahydrolase	|H2O| + |Inositol 4-phosphate| => |Phosphate| + |H+| + |L-Inositol|	3.1.3.25	R01186	0
1D-myo-Inositol 3-phosphate phosphahydrolase	|H2O| + |Inositol 3-phosphate| => |Phosphate| + |H+| + |L-Inositol|	3.1.3.25	R01187	0
Pyruvate:carbon-dioxide ligase (ADP-forming)	|ATP| + |Pyruvate| + |H2CO3| <=> |ADP| + |Phosphate| + |Oxaloacetate| + |H+|	6.4.1.1	R00344	-914.722
Heme O synthase	|H2O| + |Heme| + |Farnesyldiphosphate| <=> |PPi| + |hemeO|	2.5.1.-	None	0
cytochrome-c oxidase (H+/e- = 2)	(0.5) |O2| + (6) |H+| + (2) |Cytochrome c2+| <=> |H2O| + (4) |H+[e]| + (2) |Cytochrome c3+|	1.9.3.1,,1.9.3.1,	None	1000
Ferrocytochrome-c:oxygen oxidoreductase	|O2| + (4) |H+| + (4) |Cytochrome c2+| <=> (2) |H2O| + (4) |Cytochrome c3+|	1.9.3.1	R00081	-405.687
ATP:pantetheine-4'-phosphate adenylyltransferase	|ATP| + |Phosphopantetheine| <=> |PPi| + |Dephospho-CoA|	2.7.7.3	R03035	2.51962
UDP-N-acetylmuramoyl-L-alanyl-D-glutamate:meso-2,6-	|ATP| + |meso-2,6-Diaminopimelate| + |UDP-N-acetylmuramoyl-L-alanyl-D-glutamate| => |ADP| + |Ph	6.3.2.13	R02788	0.286573
UDPMurAc(oyl-L-Ala-D-gamma-Glu-L-Lys-D-Ala-D-Ala):undecaprenyl-	|Undecaprenylphosphate| + |UDPMurAc(oyl-L-Ala-D-gamma-Glu-L-Lys-D-Ala-D-Ala)| <=> |UMP| + |Mu	2.7.8.13	R05629	0
UDPMurAc(oyl-L-Ala-D-gamma-Glu-L-Lys-D-Ala-D-Ala):undecaprenyl-	|Undecaprenylphosphate| + |UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-6-carboxy-L-lysyl-D-alanyl- D-	2.7.8.13	R05630	0.286573
UDP-N-acetylmuramoyl-L-alanine:D-glutamate ligase(ADP-forming)	|ATP| + |D-Glutamate| + |UDP-N-acetylmuramoyl-L-alanine| => |ADP| + |Phosphate| + |H+| + |UDP-N	6.3.2.9	R02783	0.286573
Uracil hydro-lyase (adding D-ribose 5-phosphate)	|Uracil| + |ribose-5-phosphate| => |H2O| + |Pseudouridine 5'-phosphate|	4.2.1.70	R01055	0
Uracil ion-coupled transport	|H+[e]| + |Uracil[e]| <=> |H+| + |Uracil|	TC-2.A.40,2.A.40	None	-816.154
Carbamoyl-phosphate:L-aspartate carbamoyltransferase	|L-Aspartate| + |Carbamoylphosphate| => |Phosphate| + (2) |H+| + |N-Carbamoyl-L-aspartate|	2.1.3.2	R01397	0
(S)-Dihydroorotate amidohydrolase	|H2O| + |S-Dihydroorotate| <=> |H+| + |N-Carbamoyl-L-aspartate|	3.5.2.3	R01993	0
(S)-Dihydroorotate:oxygen oxidoreductase	|O2| + |S-Dihydroorotate| => |H2O2| + |Orotate|	1.3.3.1,1.3.99.11	R01867	0
dihydoorotic acid dehydrogenase (quinone8)	|S-Dihydroorotate| + |Ubiquinone-8| => |Orotate| + |Ubiquinol-8|	1.3.3.1	None	0
dihydroorotic acid (menaquinone-8)	|S-Dihydroorotate| + |Menaquinone 8| => |Orotate| + |Menaquinol 8|	1.3.3.1	None	0
Orotidine-5'-phosphate carboxy-lyase	|H+| + |Orotidylic acid| => |CO2| + |UMP|	4.1.1.23	R00965	0
Orotidine-5'-phosphate:pyrophosphate phosphoribosyltransferase	|PPi| + |Orotidylic acid| <= |PRPP| + |Orotate|	2.4.2.10	R01870	0
ATP:(d)GMP phosphotransferase	|ATP| + |dGMP| <=> |ADP| + |dGDP|	2.7.4.12,2.7.4.8	R02090	-1000
guanylate kinase (GMP:dATP)	|dATP| + |GMP| <=> |GDP| + |dADP|	2.7.4.8	None	-986.101
ATP:(d)GMP phosphotransferase	|ATP| + |GMP| <=> |ADP| + |GDP|	2.7.4.8	R00332	1000
phosphopantothenate-cysteine ligase	|CTP| + |L-Cysteine| + |4-phosphopantothenate| => |PPi| + |CMP| + |(R)-4'-Phosphopantotheno	6.3.2.5	R04231	0
N-[(R)-4'-Phosphopantothenoyl]-L-cysteine carboxy-lyase	|H+| + |(R)-4'-Phosphopantothenoyl-L-cysteine| => |CO2| + |Phosphopantetheine|	4.1.1.36	R03269	2.51962
(R)-4'-Phosphopantothenate:L-cysteine ligase	|ATP| + |L-Cysteine| + |4-phosphopantothenate| => |PPi| + |AMP| + |(R)-4'-Phosphopantotheno	6.3.2.5	R04230	2.51962
D-Ribulose-5-phosphate 3-epimerase	|D-Ribulose5-phosphate| <=> |D-Xylulose5-phosphate|	5.1.3.1	R01529	-259.655
ATP:thiamin pyrophosphotransferase	|ATP| + |Thiamin| <=> |AMP| + |TPP|	2.7.6.2	R00619	1.25981
glycerol-3-phosphate acyltransferase (C12:0)	|Glycerol-3-phosphate| + |Dodecanoyl-ACP| => |ACP| + |1-dodecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	0
glycerol-3-phosphate acyltransferase (C14:0)	|Glycerol-3-phosphate| + |Myristoyl-ACP| => |ACP| + |1-tetradecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	0
glycerol-3-phosphate acyltransferase (C14:1)	|Glycerol-3-phosphate| + |Tetradecenoyl-ACP| => |ACP| + |1-tetradec-7-enoyl-sn-glycerol 3-phosphate	2.3.1.15	None	0
glycerol-3-phosphate acyltransferase (C16:0)	|Glycerol-3-phosphate| + |Palmitoyl-ACP| => |ACP| + |1-hexadecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	0
glycerol-3-phosphate acyltransferase (C16:1)	|Glycerol-3-phosphate| + |Hexadecenoyl-ACP| => |ACP| + |1-hexadec-9-enoyl-sn-glycerol 3-phosphate	2.3.1.15	None	0
glycerol-3-phosphate acyltransferase (C18:0)	|Glycerol-3-phosphate| + |Octadecanoyl-ACP| => |ACP| + |1-octadecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	0
glycerol-3-phosphate acyltransferase (C18:1)	|Glycerol-3-phosphate| + |Octadecenoyl-ACP| => |ACP| + |1-octadec-11-enoyl-sn-glycerol 3-phosphate	2.3.1.15	None	0
palmitoyl-glycerol-3-phosphate O-acyltransferase	|Glycerol-3-phosphate| + |Palmitoyl-CoA| => |CoA| + |1-hexadecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	0
myristoyl-glycerol-3-phosphate O-acyltransferase	|Glycerol-3-phosphate| + |Myristoyl-CoA| => |CoA| + |1-tetradecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	0
isotetradecanoyl-glycerol-3-phosphate O-acyltransferase	|Glycerol-3-phosphate| + |fa1coa| => |CoA| + |1-isotetradecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	0
isopentadecanoyl-glycerol-3-phosphate O-acyltransferase	|Glycerol-3-phosphate| + |fa3coa| => |CoA| + |1-isopentadecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	0
anteisopentadecanoyl-glycerol-3-phosphate O-acyltransferase	|Glycerol-3-phosphate| + |fa4coa| => |CoA| + |1-anteisopentadecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	0
isohexadecanoyl-glycerol-3-phosphate O-acyltransferase	|Glycerol-3-phosphate| + |fa6coa| => |CoA| + |1-isohexadecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	0
stearoyl-glycerol-3-phosphate O-acyltransferase	|Glycerol-3-phosphate| + |strcoa| => |CoA| + |1-octadecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	6.59648
isoheptadecanoyl-glycerol-3-phosphate O-acyltransferase	|Glycerol-3-phosphate| + |fa11coa| => |CoA| + |1-isoheptadecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	6.59648
anteisoheptadecanoyl-glycerol-3-phosphate O-acyltransferase	|Glycerol-3-phosphate| + |fa12coa| => |CoA| + |1-anteisoheptadecanoyl-sn-glycerol 3-phosphate|	2.3.1.15	None	6.59648
2-methylpropionyl-CoA:[acyl-carrier-protein] transferase	|Isobutyryl-CoA| + |ACP| <=> |CoA| + |isobutyryl-ACP|	2.3.1.0	None	0
12-methyl-tetra-decanoyl-ACP:[acyl-carrier-protein] transferase	|CoA| + |12-methyl-tetra-decanoyl-ACP| <=> |fa4coa| + |ACP|	2.3.1.0	None	0
13-methyl-tetra-decanoyl-ACP:[acyl-carrier-protein] transferase	|CoA| + |13-methyl-tetra-decanoyl-ACP| <=> |fa3coa| + |ACP|	2.3.1.0	None	0
12-methyl-tridecanoyl-ACP:[acyl-carrier-protein] transferase	|CoA| + |12-methyl-tridecanoyl-ACP| <=> |fa1coa| + |ACP|	2.3.1.0	None	0
14-methyl-pentadecanoyl-ACP:[acyl-carrier-protein] transferase	|CoA| + |14-methyl-pentadecanoyl-ACP| <=> |fa6coa| + |ACP|	2.3.1.0	None	0
tetradecanoyl-ACP:[acyl-carrier-protein] transferase	|CoA| + |Myristoyl-ACP| <=> |Myristoyl-CoA| + |ACP|	2.3.1.1	None	0
hexadecanoyl-ACP:[acyl-carrier-protein] transferase	|CoA| + |hexadecanoyl-acp| <=> |Palmitoyl-CoA| + |ACP|	2.3.1.2	None	0
2-methylbutanoyl-CoA:[acyl-carrier-protein] transferase	|2-Methylbutyryl-CoA| + |ACP| <=> |CoA| + |2-methylbutyryl-ACP|	2.3.1.0	None	13.193
3-methylbutanoyl-CoA:[acyl-carrier-protein] transferase	|Isovaleryl-CoA| + |ACP| <=> |CoA| + |isovaleryl-ACP|	2.3.1.0	None	13.193
14-methyl-hexa-decanoyl-ACP:[acyl-carrier-protein] transferase	|CoA| + |14-methyl-hexa-decanoyl-ACP| <=> |fa12coa| + |ACP|	2.3.1.0	None	13.193
15-methyl-hexa-decanoyl-ACP:[acyl-carrier-protein] transferase	|CoA| + |15-methyl-hexa-decanoyl-ACP| <=> |fa11coa| + |ACP|	2.3.1.0	None	13.193
stearyl-ACP:[acyl-carrier-protein] transferase	|CoA| + |Octodecanoyl-ACP| <=> |strcoa| + |ACP|	2.3.1.3	None	13.193
malonyl-ACP:[acyl-carrier-protein] transferase	|Malonyl-CoA| + |ACP| <=> |CoA| + |Malonyl-acyl-carrierprotein-|	2.3.1.0,2.3.1.39,2.3.1.85,	R01626	263.859
Succinate:CoA ligase (ADP-forming)	|ATP| + |CoA| + |Succinate| <=> |ADP| + |Phosphate| + |Succinyl-CoA|	6.2.1.4,6.2.1.5	R00405	0
ATP:nucleoside-phosphate phosphotransferase	|ATP| + |UMP| <=> |ADP| + |UDP|	2.7.4.4,2.7.4.14,2.7.4.22,	R00158	-464.29
ATP:dTMP phosphotransferase	|ATP| + |dUMP| <=> |ADP| + |dUDP|	2.7.4.4,2.7.4.9	R02098	-344.252
CDP-diacylglycerol synthetase (n-C12:0)	|CTP| + |1,2-dididecanoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-dididecanoylglycerol|	2.7.7.41	None	0
CDP-diacylglycerol synthetase (n-C14:0)	|CTP| + |1,2-ditetradecanoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-ditetradecanoylglycerol|	2.7.7.41	None	0
CDP-diacylglycerol synthetase (n-C14:1)	|CTP| + |1,2-ditetradec-7-enoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-ditetradec-7-enoylglyce	2.7.7.41	None	0
CDP-diacylglycerol synthetase (n-C16:0)	|CTP| + |1,2-dihexadecanoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-dihexadecanoylglycerol|	2.7.7.41	None	0
CDP-diacylglycerol synthetase (n-C16:1)	|CTP| + |1,2-dihexadec-9-enoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-dihexadec-9-enoylglyce	2.7.7.41	None	0
CDP-diacylglycerol synthetase (n-C18:1)	|CTP| + |1,2-dioctadec-11-enoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-dioctadec-11-enoylglyc	2.7.7.41	None	0
isotetradecanoyl-phosphatidate cytidylyltransferase	|CTP| + |1,2-diisotetradecanoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-diisotetradecanoylglyce	2.7.7.41	None	0
isopentadecanoyl-phosphatidate cytidylyltransferase	|CTP| + |1,2-diisopentadecanoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-diisopentadecanoylgly	2.7.7.41	None	0
anteisopentadecanoyl-phosphatidate cytidylyltransferase	|CTP| + |1,2-dianteisopentadecanoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-dianteisopentadec	2.7.7.41	None	0
isohexadecanoyl-phosphatidate cytidylyltransferase	|CTP| + |1,2-diisohexadecanoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-diisohexadecanoylglyce	2.7.7.41	None	0
CDP-diacylglycerol synthetase (n-C18:0)	|CTP| + |1,2-dioctadecanoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-dioctadecanoylglycerol|	2.7.7.41	None	5.45019
isoheptadecanoyl-phosphatidate cytidylyltransferase	|CTP| + |1,2-diisoheptadecanoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-diisoheptadecanoylgly	2.7.7.41	None	5.45019
anteisoheptadecanoyl-phosphatidate cytidylyltransferase	|CTP| + |1,2-dianteisoheptadecanoyl-sn-glycerol 3-phosphate| <=> |PPi| + |CDP-1,2-dianteisoheptadec	2.7.7.41	None	5.45019
1-Deoxy-D-xylulose-5-phosphate isomeroreductase	|NADP| + |2-C-methyl-D-erythritol4-phosphate| <=> |NADPH| + |H+| + |1-deoxy-D-xylulose5-phospha	1.1.1.267	R05688	-33.3877
ATP:FMN adenylyltransferase	|ATP| + |FMN| => |PPi| + |FAD|	2.7.7.2	R00161	1.25981
ATP:riboflavin 5'-phosphotransferase	|ATP| + |Riboflavin| <=> |ADP| + |FMN|	2.7.1.26	R00549	1.25981
L-Aspartate-4-semialdehyde:NADP+ oxidoreductase (phosphorylating)	|NADP| + |Phosphate| + |L-Aspartate4-semialdehyde| <= |NADPH| + |4-Phospho-L-aspartate|	1.2.1.11	R02291	-0.286573
ATP:L-aspartate 4-phosphotransferase	|ATP| + |L-Aspartate| + |H+| <=> |ADP| + |4-Phospho-L-aspartate|	2.7.2.4	R00480	0.286573

FVA for reactions in the metabolic model

KeggRxnID	minFlux	maxFlux
'R00004'	999.999714	1000
'R00006'	0	0.003946539
'R00009'	0	336.3402558
'R00014'	0	0.003946539
'R00017'	142.257084	728.5980409
'R00019'	0	0
'R00026'	0	500.0005608
'R00028'	0	0
'R00036'	0	0.000105189
'R00066'	0	7.21E-05
'R00078'	0.29967889	0.299678895
'R00081'	0	428.871458
'R00082'	96.95772734	250
'R00084'	0	2.63E-05
'R00086'	-1000	-999.9993769
'R00103'	0	0.000387478
'R00104'	2.397076875	2.397431156
'R00114'	0	1000
'R00124'	-1000	-999.9999579
'R00127'	-1000	-999.9997576
'R00130'	0	0.000593299
'R00131'	1.19871556	1.202267445
'R00132'	0	0
'R00137'	-0.000263363	0.000387478
'R00156'	0	1000
'R00158'	-199.3217022	-199.3183251
'R00160'	0	0.000387478
'R00161'	1.19871556	1.199103038
'R00177'	3.613818253	3.614067272
'R00178'	0	7.64E-05
'R00183'	0	0.001109964
'R00185'	0	0.001420754
'R00189'	0	0.000499328
'R00190'	-25.07093539	-25.05739983
'R00194'	0	0.00053623
'R00200'	-1000	1000
'R00209'	0	0.00126853
'R00212'	248.3695159	248.3859488
'R00214'	999.9971201	1000
'R00217'	0	0.002879907
'R00220'	385.3870772	1000
'R00226'	0	0.003946539
'R00228'	999.9971201	1000
'R00230'	986.5103309	986.5121507
'R00235'	0	0.001024582
'R00405'	232.231738	232.2482058
'R02164'	-232.2482058	-232.231738
'R00414'	0	1000
'R00416'	598.2095668	598.2150202

KeggRxnID	minFlux	maxFlux
'R00236'	-1000	-999.9989754
'R00238'	-234.6456369	-234.6291691
'R00239'	0	0.000623138
'R00243'	-1000	1000
'R00245'	-1000	998.8019076
'R00248'	-1000	1000
'R00253'	0	0.000623138
'R00254'	0	0
'R00256'	-1000	0.002746303
'R00257'	0	0.000499328
'R00259'	0	0.000623138
'R00260'	-207.1936223	-61.11512929
'R00267'	0	1000
'R00268'	-49.72919387	950.2779099
'R00272'	0	0.000352836
'R00274'	0	0.001869413
'R00276'	0	0
'R00286'	999.9984096	1000
'R00289'	28.7509782	1000
'R00291'	28.75176503	28.7536478
'R00306'	-500.0005608	0
'R00307'	-1000	0.001121648
'R00310'	0	0
'R00315'	-986.5121507	-986.5103309
'R00316'	-1000	-999.9989754
'R00317'	0	0.000623138
'R00321'	0	0
'R00330'	-1000	1000
'R00332'	-5.853007475	-5.849880987
'R00335'	0	0.000623138
'R00336'	0	0.000238916
'R00341'	0	0.002879907
'R00342'	999.9971201	1000
'R00344'	0	0
'R00351'	-1000	-999.9893443
'R00355'	-1000	988.1869725
'R00357'	11.81302751	1000
'R00365'	0	0
'R00366'	0	0
'R00375'	333.9330997	333.9381738
'R00376'	-0.939175061	-0.933102309
'R00377'	-331.8040888	-331.7990146
'R00378'	-1.198715578	-1.197581979
'R00396'	868.3565269	1000
'R00399'	846.4875875	846.510872
'R00582'	0	0.000488792
'R00586'	-1000	-999.9993769
'R00590'	385.3870772	1000
'R00615'	0	0.000623138

FVA for reactions in the metabolic model

'R00420'	-401.7904332	598.2150202
'R00424'	0	0
'R00425'	0	7.21E-05
'R00428'	0	0
'R00429'	0	0.000238916
'R00430'	0	1000
'R00435'	999.9996624	1000
'R00437'	-1000	-999.9996624
'R00438'	797.5301591	1000
'R00439'	5.846934723	1000
'R00440'	-596.7080432	599.914863
'R00441'	-1000	-5.846934723
'R00442'	-599.914863	596.7080432
'R00443'	-1000	-797.5301591
'R00451'	13.48887386	13.48887407
'R00465'	-988.1869725	1000
'R00472'	0	0
'R00475'	11.81302751	1000
'R00479'	0	0
'R00480'	13.48887386	13.49074327
'R00481'	999.9985979	1000
'R00483'	0	806.2994178
'R00485'	0	0.000387478
'R00489'	0	0.000880633
'R00494'	410.430954	410.4449052
'R00502'	-0.001590396	971.248235
'R00511'	0	0.000623138
'R00512'	-3.209086942	-3.206543585
'R00513'	-1000	1000
'R00516'	-1000	1000
'R00517'	-1000	1000
'R00519'	0	0.00126853
'R00522'	0	0
'R00549'	1.19871556	1.198715578
'R00551'	0	0.003551885
'R00566'	1.19871556	1.199036513
'R00570'	-599.914863	1000
'R00571'	0	0.000859327
'R00572'	0	1000
'R00573'	0	0.000859327
'R00575'	0	0
'R00578'	-1000	-193.7005822
'R00770'	0	1000
'R00771'	-1000	0
'R00772'	-1000	0.000859327
'R00782'	-1000	0.000623138
'R00783'	0	765.2113633
'R00784'	387.8309094	1000
'R00790'	0	0
'R00802'	-971.2479546	-971.2463522
'R00616'	0	0
'R00617'	0	0.000623138
'R00619'	1.198623239	1.198715578
'R00621'	0	0.000813409
'R00650'	0	0.00020143
'R00651'	0	0.000688731
'R00654'	0	0.000688731
'R00658'	0	0.002267161
'R00659'	0	1000
'R00660'	0	4.00E-05
'R00667'	-1.199036513	-1.198092422
'R00668'	0	0
'R00669'	0	0.000623138
'R00674'	0	0
'R00688'	-1000	1000
'R00691'	0	1000
'R00694'	-1000	1000
'R00703'	0	0
'R00705'	-474.4942188	525.5093331
'R00706'	0	1000
'R00707'	-1000	1000
'R00708'	-1000	1000
'R00710'	-1000	-999.9992048
'R00713'	0	0.000352836
'R00714'	0	0.000352836
'R00717'	-1000	988.1869725
'R00722'	-1000	1000
'R00724'	0	1000
'R00732'	0	1000
'R00734'	-1000	7.059484123
'R00740'	0	0
'R00742'	0	0
'R00750'	842.2311727	842.2392569
'R00754'	999.9971201	1000
'R00756'	-1000	0.00140206
'R00758'	-1000	1000
'R00760'	-1000	971.248235
'R00762'	0	0.000623138
'R00765'	-598.2150202	401.7904332
'R00767'	0	1000
'R00768'	0	1000
'R00769'	0	1000
'R00946'	0	0.000200294
'R00948'	0	0
'R00959'	-942.4964532	57.5072956
'R00962'	-1000	1000
'R00963'	0	0.000623138
'R00964'	-1000	1000
'R00965'	1.198323808	1.202267444
'R00966'	194.6385842	194.643246

FVA for reactions in the metabolic model

'R00803'	971.2463522	971.2479546
'R00816'	842.2311727	842.2392569
'R00830'	0	0.000206305
'R00833'	231.0330224	1000
'R00835'	0	1000
'R00837'	0	0.000245635
'R00838'	0	0
'R00839'	0	0
'R00842'	-1000	1000
'R00844'	-1000	1000
'R00847'	-1000	-999.9991928
'R00848'	-1000	-999.9973828
'R00856'	0	0
'R00867'	0	1000
'R00875'	0	0
'R00878'	-1000	1000
'R00885'	999.9991407	1000
'R00895'	405.8433301	405.8572813
'R00896'	0	0.002267161
'R00897'	0	0.000623138
'R00899'	410.430954	410.4449052
'R00904'	0	0
'R00905'	525.5057812	525.5093331
'R00908'	525.5057812	525.5093331
'R00921'	0	0.001024582
'R00922'	0	0
'R00924'	0	0.001090491
'R00925'	0	0.001024582
'R00926'	0	0.001024582
'R00927'	0	0
'R00935'	0	0.00042843
'R00936'	0	406.1489988
'R00937'	0	406.1489988
'R00939'	-1000	-593.8510012
'R00940'	0	406.1489988
'R00943'	593.8510012	593.8616568
'R00944'	0	0.000623138
'R00945'	-393.0996727	-393.0910493
'R01092'	971.2463522	971.248235
'R01103'	0	0
'R01104'	-1000	-999.9988784
'R01106'	428.1713894	1000
'R01126'	0	0.001109964
'R01127'	-0.000178786	0
'R01130'	0	0.000302717
'R01132'	-170.7834518	-170.7699162
'R01135'	170.7699162	170.7834518
'R01137'	-1000	1000
'R01138'	-1000	1000
'R01148'	0	131.6434731
'R00967'	-1000	1000
'R00968'	-1000	1000
'R00970'	-1000	1000
'R00978'	-525.5093331	-525.5057812
'R00985'	0	0.000942978
'R00986'	0	0.000942978
'R00994'	0	0
'R00996'	0	0.001090491
'R00999'	0	0.000688731
'R01000'	0	0
'R01001'	0	0.000688731
'R01010'	999.9993769	1000
'R01015'	999.9980626	1000
'R01016'	0	0.001109964
'R01030'	0	0
'R01034'	999.9993769	1000
'R01049'	0	0.000595288
'R01051'	0	0.001522236
'R01054'	0	0.000154519
'R01055'	0	0
'R01056'	684.4696645	684.4720324
'R01057'	526.7044968	526.7080487
'R01061'	-0.000623138	1000
'R01063'	-0.000623138	1000
'R01066'	157.7607431	157.7688273
'R01067'	-842.2365106	1000
'R01068'	-0.002267161	0.00140206
'R01070'	999.9977328	1000
'R01071'	0	0.000384681
'R01072'	0	0
'R01073'	-0.000942978	0
'R01074'	0	0
'R00955'	-1000	-28.7509782
'R01082'	-1000	-999.9993382
'R01083'	170.7699162	170.7834518
'R01086'	1.198323808	1.202267444
'R01088'	-14.77834643	985.2245337
'R01090'	0	1000
'R01288'	0	1000
'R01291'	0	0.00053623
'R01324'	-0.010655655	1000
'R01325'	-1000	0.010655655
'R01329'	0	0
'R01353'	0	0.001024582
'R01354'	0	0.001024582
'R01357'	0	0
'R01360'	0	0
'R01373'	-6.17179754	1000
'R01385'	999.9984096	1000
'R01388'	-1000	1000

FVA for reactions in the metabolic model

'R01150'	0	5.48E-05
'R01155'	0	0.000320953
'R01157'	1.19871556	1.199036513
'R01158'	0	0.000384681
'R01163'	0	0.000384681
'R01166'	0	0
'R01172'	0	1000
'R01173'	0	1000
'R01174'	0	0.001090491
'R01175'	285.7139118	285.7152204
'R01177'	-285.7142857	-285.7139118
'R01185'	0	0
'R01194'	0	0
'R01209'	0	0.003946539
'R01213'	14.7754663	14.78061336
'R01214'	-1000	-14.77439967
'R01216'	0	0.001479952
'R01220'	-393.0996727	-393.089017
'R01221'	0	0.000661842
'R01224'	-1000	0.001590396
'R01226'	0	0.001479952
'R01227'	0	0.001109964
'R01229'	-0.002267161	0.001109964
'R01230'	-1000	0.000302717
'R01231'	0	1000
'R01244'	0	0.002267161
'R01248'	-7.085120823	1000
'R01251'	-7.085120823	1000
'R01257'	-1000	-999.9973828
'R01268'	0	0.000263363
'R01279'	285.7139118	285.7142857
'R01285'	0	1000
'R01286'	0	1000
'R01287'	-1000	0.000623138
'R01655'	-393.0996727	-393.089017
'R01658'	0	0
'R01652'	14.7754663	14.78061336
'R01664'	0	0.000623138
'R01665'	-331.8027932	-331.7990146
'R01666'	0	0.000623138
'R01663'	330.8636452	330.8674238
'R01698'	0	0
'R01714'	0	0.000119191
'R01715'	-1.199658538	-1.197690977
'R01716'	0	0
'R01717'	0	0.000250133
'R01718'	0	0
'R01724'	-1000	-999.9985979
'R01728'	-1000	7.059484123
'R01731'	0	1000
'R01392'	-1000	1000
'R01394'	0	0
'R01397'	1.198323808	1.202267444
'R01398'	1.198323808	1.202267444
'R01401'	0	7.64E-05
'R01432'	0	0
'R01433'	0	0
'R01434'	0	985.2256003
'R01466'	0	0.001869413
'R01492'	0	0.000109649
'R01504'	0	0
'R01505'	0	0
'R01512'	-1000	-999.9993769
'R01513'	0	0.000488792
'R01514'	-1000	-999.9993769
'R01515'	0	0.000623138
'R01518'	0	0.000623138
'R01526'	0	0
'R01529'	-315.5303355	-315.5269789
'R01535'	0	0.000125716
'R01547'	333.9330997	333.9381738
'R01548'	-1000	1000
'R01549'	-1000	1000
'R01561'	-0.001420754	355.9355405
'R01567'	0	0.000623138
'R01569'	0	0.000623138
'R01579'	61.11512929	61.13033382
'R01582'	0	14.43094
'R01600'	0	1000
'R01602'	-1000	1000
'R01639'	0	0
'R01641'	157.7634894	157.7651677
'R01648'	0	0.000320953
'R01654'	200.7618062	200.7622588
'R01899'	-49.72919387	950.2779099
'R01900'	-950.2814618	49.72919387
'R01902'	0	0
'R01920'	0	3.82E-05
'R01954'	1.198323808	1.202267444
'R01967'	-1000	-999.9977328
'R01968'	-1000	-999.9991407
'R01969'	-0.002267161	0.000859327
'R01975'	-1000	1000
'R01976'	-1000	1000
'R01986'	0	1000
'R01992'	0	0
'R01993'	-1.202267444	-1.198323808
'R02003'	0	0
'R02016'	0	0.003805591
'R02017'	0	0.000310118

FVA for reactions in the metabolic model

'R01736'	0	0	'R02019'	0	0.003805591
'R01751'	0	0	'R02024'	0	0.003805591
'R01752'	-1000	-999.9977328	'R02047'	0	0
'R01761'	0	0	'R02059'	0	0
'R01762'	0	0	'R02060'	598.2095668	598.2150202
'R01771'	0	0.001869413	'R02061'	0	0
'R01773'	-1000	1000	'R02071'	999.9977328	1000
'R01775'	-1000	1000	'R02088'	-335.1368893	-335.1318152
'R01777'	0	0.000688731	'R02089'	0	0.000623138
'R01785'	0	0	'R02090'	-0.936228797	-0.933102309
'R01786'	0	1000	'R02091'	-1000	1000
'R01791'	0	0	'R02093'	-1000	1000
'R01800'	0	0.000276196	'R02094'	-1.198715578	-1.197581979
'R01804'	0	0	'R02096'	-1000	1000
'R01818'	999.9991407	1000	'R02097'	-1000	1000
'R01819'	-1000	0.000859327	'R02098'	0	0
'R01821'	0	0	'R02099'	-330.8674238	-330.8636452
'R01826'	0	0.000119191	'R02101'	0	0.00113358
'R01827'	-1000	-157.7634894	'R02102'	0	0.000623138
'R01830'	-842.2365106	1000	'R02108'	0	0
'R01843'	0	0.000623138	'R02110'	999.9991988	1000
'R01845'	0	0.000623138	'R02111'	999.9991988	1000
'R01857'	-1000	1000	'R02142'	0	0.001109964
'R01858'	0	1000	'R02147'	0	0.001109964
'R01859'	0	0	'R02163'	0	1000
'R01863'	170.7699162	526.7080487	'R02165'	0	448.4536744
'R01867'	-1.202267444	-1.198323808	'R02196'	-1000	1000
'R01870'	-1.202267444	-1.198323808	'R02199'	-1000	1000
'R01878'	0	0.002267161	'R02235'	0	406.1489988
'R01880'	0	1000	'R02236'	-1000	-593.8510012
'R02237'	0	0	'R02662'	0	0.00126853
'R02269'	525.5057812	525.5093331	'R02663'	0	0.000634265
'R02272'	0	0.000210378	'R02703'	-1000	1000
'R02282'	-0.000623138	0.000623138	'R02704'	0	0
'R02283'	0	0.000623138	'R02707'	0	1000
'R02291'	13.48887386	13.49074327	'R02719'	0	0.001109964
'R02292'	986.5092567	986.5111261	'R02722'	0	0.000942978
'R02294'	0	0.000263363	'R02733'	13.48887386	13.48891385
'R02295'	0	0.000310661	'R02735'	13.48887386	13.48891385
'R02296'	0	0	'R02736'	0	1000
'R02297'	0	0.001109964	'R02738'	0	0.000354244
'R02300'	0	0	'R02739'	-1000	1000
'R02301'	0	0	'R02740'	-1000	1000
'R02320'	0	0.000354244	'R02748'	644.0644595	1000
'R02323'	0	0.000263363	'R02749'	999.9984778	1000
'R02326'	-1000	1000	'R02750'	842.2311727	842.2392569
'R02327'	-1000	1000	'R02762'	0	0.00126853
'R02331'	-1000	1000	'R02765'	-1000	-231.0330224
'R02332'	-1000	1000	'R02780'	0	0.000245635
'R02371'	-1000	1000	'R02783'	0	4.00E-05

FVA for reactions in the metabolic model

'R02372'	-1000	1000	'R02788'	0	4.00E-05
'R02412'	0	0.000119191	'R02869'	0	3.82E-05
'R02413'	-0.000119191	0	'R02886'	0	0
'R02421'	0	0	'R02887'	0	0
'R02439'	0	0	'R02926'	0	0
'R02472'	-0.000593299	0	'R02964'	0	0
'R02473'	0	0.000593299	'R02971'	0	0
'R02484'	330.8636452	330.8674238	'R03004'	0	0.000387478
'R02485'	0	0.000623138	'R03005'	0	0.000499328
'R02508'	0	1000	'R03012'	0	0.000384681
'R02527'	0	0	'R03013'	0	0.000384681
'R02528'	0	0.001109964	'R03018'	0	0.000593299
'R02530'	0	0	'R03035'	0	0.000593299
'R02545'	0	0	'R03036'	0	0.000387478
'R02549'	-1000	0.000320953	'R03037'	0	0
'R02557'	-330.867197	25.07093539	'R03050'	-0.003946539	0.003946539
'R02565'	0	0	'R03051'	0	0.003946539
'R02568'	-1000	-999.9977328	'R03066'	-0.000387478	0
'R02569'	0	0.00126853	'R03067'	0	0.000387478
'R02570'	0	0.000813409	'R03083'	0	0.000119191
'R02601'	842.2311727	842.2392569	'R03084'	0	0.000119191
'R02602'	0	0.00126853	'R03104'	0	0.001590396
'R02604'	842.2311727	842.2392569	'R03105'	405.8433301	405.8572813
'R02630'	0	0.000269218	'R03165'	0	1.38E-05
'R02631'	0	0.000240643	'R01931'	405.8433301	405.8572813
'R02649'	0	0.000623138	'R03174'	0	0.00126853
'R03175'	0	0.00126853	'R03777'	285.7139118	285.7142857
'R00858'	0	0.001869413	'R03778'	-285.7142857	-285.7139118
'R03191'	0	4.00E-05	'R03815'	0	0.000661842
'R03192'	0	4.00E-05	'R03857'	285.7139118	285.7142857
'R03193'	0	4.00E-05	'R03858'	-285.7142857	-285.7139118
'R03194'	0	1.38E-05	'R03869'	0	0.000623138
'R03197'	0	0	'R03920'	0	0.000623138
'R03217'	-0.000623138	1000	'R03921'	0	0
'R03222'	0	0	'R03948'	0	1.38E-05
'R03223'	0	9.23E-05	'R03966'	0	0.00126853
'R03232'	0	0.000306373	'R03968'	14.7754663	14.78061336
'R03243'	0	0.000384681	'R03990'	285.7139118	285.7142857
'R03260'	0	1000	'R03991'	-285.7142857	-285.7139118
'R03269'	0	0.000593299	'R04001'	14.7754663	14.78061336
'R03270'	0	0.00126853	'R04030'	0	0.000250133
'R03313'	0	0.000623138	'R04031'	0	0.000250133
'R03316'	0	0.000813409	'R04035'	0	0.000384681
'R03317'	598.2095668	598.2150202	'R04037'	0	0.000384681
'R03321'	0	1000	'R04095'	0	0.000736324
'R03346'	0	0.000310661	'R04097'	0	0.002267161
'R03314'	-1000	998.8019076	'R04098'	0	0.002267161
'R03348'	999.9985979	1000	'R04111'	0	0
'R03381'	0	0.000623138	'R04112'	0	0
'R03409'	0	0.000238916	'R04109'	0	0.000210378

FVA for reactions in the metabolic model

'R03425'	0	0.000661842	'R04125'	0	0.000661842
'R03443'	0	0.000623138	'R04143'	0	7.64E-05
'R03457'	0	0.000384681	'R04144'	0	0
'R03458'	0	7.21E-05	'R04148'	0	0.000109649
'R03459'	0	7.21E-05	'R04150'	0	0
'R03460'	0	0.000119191	'R04170'	-285.7142857	-285.7139118
'R03471'	0	9.23E-05	'R04173'	0	0.000488792
'R03472'	0	0	'R04188'	0	0
'R03503'	0	0.000387478	'R04198'	0	986.5111261
'R03504'	0	0	'R04199'	13.48887386	1000
'R03508'	0	0.000942978	'R04203'	0	0
'R03509'	0	0.000942978	'R04208'	0	0
'R03544'	0	1000	'R04209'	0	0
'R03545'	0	1000	'R04230'	0	0.000593299
'R03560'	0	1000	'R04231'	0	0.000172087
'R03562'	0	1000	'R04241'	0	0
'R03566'	0	1000	'R04292'	999.9985979	1000
'R03596'	0	0.000327147	'R04325'	0	0
'R03599'	0	0.000467353	'R04326'	0	0
'R03601'	0	0.000467353	'R04364'	13.48887386	13.48891385
'R03608'	-1000	0	'R04378'	0	0.000384681
'R03616'	999.9988784	1000	'R04391'	0	0
'R04394'	0	0.000314512	'R04754'	285.7139118	285.7142857
'R04405'	0	0.000219704	'R04770'	0	0.000327147
'R04426'	14.7754663	14.78061336	'R04773'	0	0.000172701
'R04420'	0	7.64E-05	'R04779'	0	0.002267161
'R07392'	0	7.64E-05	'R04780'	0	0.000623138
'R07393'	0	7.64E-05	'R10404'	0	0.00053623
'R04439'	-1000	1000	'R04930'	0	0.000688731
'R04440'	-1000	1000	'R04937'	0	0
'R04441'	0	0.003946539	'R04938'	0	0
'R04448'	0	9.23E-05	'R04941'	0	0.000327147
'R07394'	0	7.64E-05	'R04944'	0	0
'R07364'	0	7.64E-05	'R04945'	0	0.000688731
'R04457'	0	0.000144132	'R04946'	0	0.000688731
'R04463'	0	0	'R03166'	0	2.63E-05
'R04467'	13.48887386	13.48891385	'R04972'	0	2.63E-05
'R04509'	0	9.23E-05	'R04993'	0	0
'R04519'	0	4.00E-05	'R05032'	0	4.00E-05
'R07456'	999.9972537	1000	'R05046'	0	0
'R04558'	0	0.000384681	'R05048'	0	0
'R04559'	0	0	'R05066'	0	0
'R04560'	0	0.000178786	'R05068'	0	0.00075572
'R02133'	0	9.23E-05	'R05069'	0	0.00075572
'R04573'	0	5.48E-05	'R05070'	0	0.00075572
'R04591'	0	0	'R05071'	-1000	1000
'R04594'	0	0.000109649	'R05132'	0	0.000314512
'R04617'	0	4.00E-05	'R05133'	0	0.000314512
'R04620'	0	0	'R05134'	0	0.000314512
'R04639'	0	0	'R05149'	0	1.38E-05

FVA for reactions in the metabolic model

'R04640'	0	0.000384681	'R05150'	0	1.38E-05
'R04672'	-0.003946539	0.003946539	'R05177'	0	1.38E-05
'R04673'	-0.003946539	0.00075572	'R05180'	0	1.38E-05
'R04698'	0	0	'R05181'	0	1.38E-05
'R04737'	285.7139118	285.7142857	'R05197'	200.7618062	200.7622588
'R04738'	-285.7142857	-285.7139118	'R05217'	0	1.38E-05
'R04739'	285.7139118	285.7142857	'R05220'	0	1.38E-05
'R04740'	-285.7142857	-285.7139118	'R05221'	0	0.000109649
'R04741'	285.7139118	285.7142857	'R05222'	0	0.000109649
'R04742'	-285.7142857	-285.7139118	'R05223'	0	0.000109649
'R04743'	285.7139118	285.7142857	'R05224'	0	1.38E-05
'R04744'	-285.7142857	-285.7139118	'R05225'	0	1.38E-05
'R04745'	285.7139118	285.7142857	'R05219'	0	1.38E-05
'R04746'	-285.7142857	-285.7139118	'R05227'	0	1.38E-05
'R04747'	-285.7142857	-285.7139118	'R05218'	0	6.89E-06
'R04748'	285.7139118	285.7142857	'R05332'	598.2095668	598.2150202
'R04749'	-285.7142857	-285.7139118	'R05338'	0	0.001479952
'R04751'	285.7139118	285.7142857	'R05339'	0	0.001479952
'R05351'	0	0	'R06178'	0	5.48E-05
'R05553'	0	0	'R06180'	0	0
'R05555'	0	0	'R06200'	0	0
'R05566'	0	0	'R06203'	0	0
'R05576'	285.7139118	285.7152204	'R06447'	0	0
'R05578'	0	0.000210378	'R06529'	0	1.30E-05
'R05595'	285.7139118	285.7152204	'R06530'	0	1.30E-05
'R05611'	0	0	'R06558'	0	0.000109649
'R05612'	0	0	'R06861'	0	0
'R05613'	0	0	'R06863'	0	0
'R05617'	0	0	'R06895'	0	0
'R05627'	0	0.000125716	'R06987'	0	0.001090491
'R05629'	0	5.48E-05	'R07168'	0	1000
'R05630'	0	4.00E-05	'R07219'	0	1000
'R05633'	0	0	'R07237'	0	1.38E-05
'R05634'	0	0	'R07238'	0	1.38E-05
'R05636'	0	0	'R07263'	0	0.000250133
'R05637'	0	0	'R07268'	0	0.000109649
'R05662'	0	5.48E-05	'R07269'	0	0
'R05688'	0	0	'R07280'	0	7.21E-05
'R05705'	0	1000	'R07281'	0	0.000144132
'R05706'	0	1000	'R07302'	0	1.38E-05
'R05721'	0	336.3402558	'R07343'	0	0
'R05724'	0	1000	'R07396'	0	7.64E-05
'R05725'	-500	882.6056816	'R07405'	0	0
'R05850'	0	0	'R07404'	0	0
'R05861'	0	131.6434731	'R07411'	0	0
'R05884'	0	0	'R07463'	0	0.010655655
'R06034'	0	0	'R07475'	0	0
'R06070'	971.2463522	971.2479546	'R07599'	0	0.00126853
'R06077'	0	0	'R07600'	0	0.00126853
'R06080'	0	0	'R07601'	0	0.002267161

FVA for reactions in the metabolic model

'R06084'	0	0
'R06087'	0	0.000623138
'R06088'	971.2463522	971.2479546
'R06091'	0	0.00050984
'R06092'	0	0
'R06096'	0	1000
'R06100'	0	1000
'R06101'	0	1000
'R06102'	0	0
'R06112'	0	0
'R06113'	0	0
'R06115'	0	0
'R06142'	0	0
'R06152'	0	1000
'R08575'	0	842.2365106
'R08555'	0	0
'R08632'	0	0.000688731
'R08635'	0	0.000688731
'R08637'	0	0.000688731
'R08639'	0	1000
'R08648'	0	0.003946539
'R08657'	0	1000
'R08689'	0	0
'R08698'	385.3870772	1000
'R08748'	0	0
'R08749'	0	0
'R08750'	0	0
'R08751'	0	0
'R08752'	0	0
'R08753'	0	0
'R08856'	0	0
'R09247'	0	0
'R09365'	0	0.000327147
'R09372'	0	0
'R01624'	0	0
'R04355'	-1000	1000
'R04533'	-1000	1000
'R01626'	0	0
'R04428'	0	0
'R04429'	0	0
'R04952'	0	0
'R04953'	0	0
'R04954'	311.6450155	1000
'R04955'	-1000	-311.6450155
'R04957'	0	0
'R04536'	0	0
'R04537'	0	0
'R04958'	0	0
'R04960'	0	0
'R04534'	0	0
'R07602'	0	0.002267161
'R07603'	0	0.00126853
'R07604'	0	0.00126853
'R07618'	0	0.002267161
'R07641'	0	0
'R08165'	0	0.000250133
'R08166'	0	0.000250133
'R08209'	0	0
'R08210'	0	1000
'R08549'	0	0.000813409
'R08557'	0	0
'R08558'	0	0
'R08559'	0	0
'R08572'	0	0.002267161
'R04968'	0	0
'R04543'	-1000	-999.9981306
'R04544'	999.9981306	1000
'R04969'	-1000	-999.9981306
'R07762'	0	0
'R07763'	0	0
'R07764'	0	0
'R07765'	0	0
'R01688'	0	0.001090491
'RNGAM'	-1000	-999.9993769
'R10619'	-1000	-999.9996458
'R10305'	-1000	-999.9993769
'R10147'	999.9981306	1000

FVA for reactions in the metabolic model

'R04535'	311.6450155	1000
'R04961'	-1000	-311.6450155
'R04963'	0	0
'R04964'	0	0
'R04965'	0	0
'R04724'	0	0
'R04726'	0	0
'R04566'	-1000	-999.9981306
'R04568'	999.9981306	1000
'R04966'	-1000	-999.9981306